PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/71, C12N 15/85, 5/10, A61K 38/17, 48/00, 39/395, G01N 33/68

A1

(11) International Publication Number: WO 00/20579

(43) International Publication Date: 13 April 2000 (13.04.00)

(21) International Application Number: PCT/CA99/00912

(22) International Filing Date: 1 October 1999 (01.10.99)

(30) Priority Data:
09/165,192 2 October 1998 (02.10.98) US

(71) Applicant *(for all designated States except US)*: McMASTER UNIVERSITY [CA/CA]; Health Science Centre, Room 3N43, 1200 Main Street, Hamilton, Ontario L8N 3Z5 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MULLER, William, J. [CA/CA]; #65 7 Davidson Boulevard, Dundas, Ontario L9H 6Y7 (CA). SIEGEL, Peter, M. [CA/CA]; Apartment 607, 200 Main Street West, Hamilton, Ontario L8S 4M8 (CA).

(74) Agent: BERESKIN & PARR; 40th floor, 40 King Street West, Toronto, Ontario M5H 3Y2 (CA).

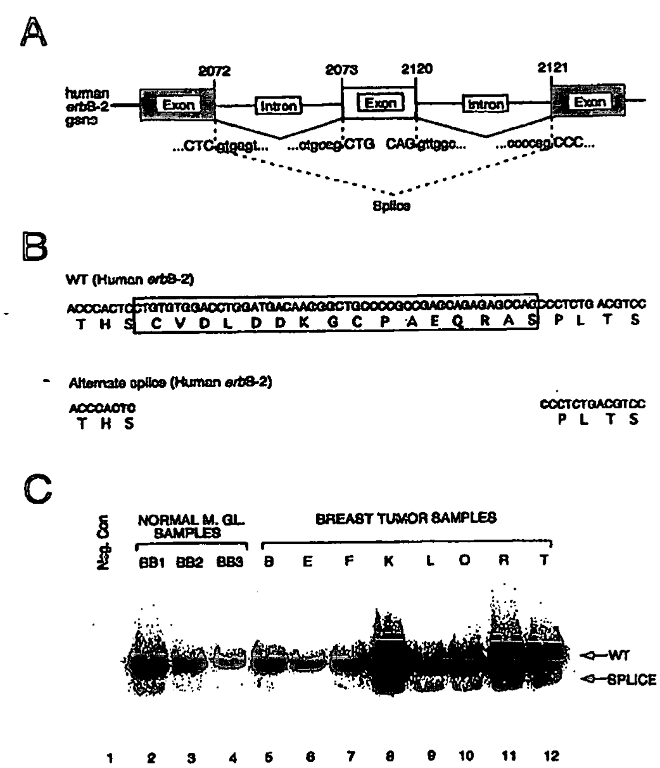
(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SPLICED FORM OF *ERB*B–2/NEU ONCOGENE

(57) Abstract

A novel spliced form of *erb*B–2 oncogene that encodes a constitutively tyrosine phosphorylated *erb*B–2 receptor is described. The protein has an in frame deletion of 16 amino acids, two of which are conserved cysteine residues. The protein may be useful in developing remedies to control cell transformation such as in breast cancer.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Title: **Spliced Form of *Erb*B-2/Neu Oncogene**

FIELD OF THE INVENTION

The invention relates to a novel spliced form of gene that encodes a constitutively tyrosine phosphorylated *erb*B-2 receptor. The invention includes nucleic acid molecules encoding the protein and truncations, analogs, and homologs of the protein; and uses of the protein and nucleic acid molecules in controlling cell transformation and cancer.

BACKGROUND OF THE INVENTION

Activation of growth factor receptor tyrosine kinases has been implicated in the malignant progression of a number of different tumor types. For example, elevated expression of the *erb*B-2 oncogene has been observed in a significant proportion of human breast and ovarian cancers (Slamon et at, 1987; 1989). Moreover, the extent of overexpression of *erb*B-2 in lymph node positive breast cancer patients has been inversely correlated with patient survival (reviewed in Hynes and Stern, 1994). Direct evidence for the importance of this gene in the induction of mammary tumors has been obtained from studies of transgenic mice that express either an activated or wild-type neu cDNA in the mammary epithelium. Mammary epithelial expression of the activated version of neu results in the rapid induction of multifocal mammary tumors in every female transgenic animal (Muller *et al.*, 1988; Guy *et al.*, 1996). In contrast, mammary epithelial expression of the neu proto-oncogene leads to development of focal mammary tumors which develop after a longer latency period (Guy *et al.*, 1992). Consistent with clinical observations indicating a negative correlation of *erb*B-2 amplification with patient survival, these wild-type neu transgenic strains frequently developed secondary metastatic lesions in the lung (Guy *et al.*, 1992).

Biochemical and genetic analyses of tumor progression in these transgenic strains has revealed that tumor progression in transgenic mice expressing the neu proto-oncogene in the mammary epithelium frequently involves the acquisition of activating mutations in the neu transgene (Siegel *et al.*, 1994). Sequence analyses of these somatic mutations revealed that these alterations map to a cysteine rich region in Neu that is highly conserved among members of the epidermal growth factor receptor family. Indeed, it has been demonstrated that these alterations result in the oncogenic activation of Neu (Siegel *et al.*, 1994; Siegel and Muller, 1996). Significantly, these sequence alterations appear to affect the balance of cysteines within this region through the deletion or insertion of a cysteine residue. It has been further demonstrated that these alterations in Neu lead to receptor dimerization and activation through the formation of intermolecular cysteine disulfide bonds (Siegel and Muller, 1996). Interestingly, oncogenic activation of these altered Neu molecules is dependent on the formation of intermolecular disulfide bonds since the addition of exogenous reducing agents can ablate the in vitro transforming potential of these activated neu alleles (Siegel and Muller, 1996).

Although these observations sugg st that activation of the Neu receptor in this transgenic mouse model is an important step, the relevance to clinical human breast cancer is unclear.

SUMMARY OF THE INVENTION

The present inventors have identified a spliced form of mRNA which translates into an *erb*B-2 receptor which has increased catalytic activity capable of transforming Rat-1 fibroblasts. The increased catalytic activity is reflected by increased tyrosine phosphorylation of the receptor. The increase in catalytic activity of this receptor may constitute a significant causative factor in the genesis of human breast tumors.

The present invention therefore provides a purified and isolated nucleic acid molecule (SPLICE) comprising a sequence encoding a constitutively tyrosine phosphorylated *erb*B-2 receptor, which is the protein of the invention, having increased catalytic activity and capable of transforming Rat-1 fibroblasts and which may generally be referred to as SPLICE *erb*B-2 protein.

In an embodiment of the invention, the purified and isolated nucleic acid molecule comprises: (i) a nucleic acid sequence encoding a SPLICE *erb*B-2 protein having the amino acid sequence as shown in Figure 2 and SEQ ID NO:2; and, (ii) nucleic acid sequences complementary to (i).

In a preferred embodiment of the invention, the purified and isolated nucleic acid molecule comprises:

(i) a nucleic acid sequence encoding a SPLICE *erb*B-2 protein having the nucleic acid sequence as shown in Figure 1 or SEQ ID NO:1, wherein T can also be U;

(ii) a nucleic acid sequence complementary to (i), preferably complementary to the full length nucleic acid sequence shown in Figure 1;

(iii) a nucleic acid molecule differing from any of the nucleic acids of (i) and (ii) in codon sequences due to the degeneracy of the genetic code.

The invention also contemplates an isolated and purified (a) nucleic acid molecule comprising a sequence encoding a truncation of SPLICE *erb*B-2, an analog, homolog or truncation thereof; (b) nucleic acid molecule comprising a sequence which hybridizes under high stringency conditions to the nucleic acid encoded by a SPLICE *erb*B-2 protein having the amino acid sequence as shown in Figure 2 or SEQ ID NO:2, wherein T can also be U, or complementary sequences thereto, or by a related protein; and (c) nucleic acid molecule comprising a sequence which hybridizes under high stringency conditions to the nucleic acid sequence as shown in Figure 1, wherein T can also be U, or complementary sequences thereto.

The invention further contemplates a purified and isolated double stranded nucleic acid molecule containing a nucleic acid molecule of the invention, hydrogen bonded to a complementary nucleic acid base sequence.

The nucleic acid molecules of the invention may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Accordingly, recombinant expression vectors adapted for transformation of a host cell, preferably mammary cells, may be constructed which comprise a nucleic acid molecule of the invention and one or more transcription and translation elements operatively linked to the nucleic acid molecule.

The recombinant expression vector can be used to prepare transformed host cells expressing a SPLICE *erb*B-2 protein of the invention or a related protein. Therefore, the invention further provides host cells containing a recombinant molecule of the invention. The invention also contemplates transgenic non-human mammals whose germ cells and somatic cells contain a recombinant molecule comprising a nucleic acid molecule of the invention which encodes a SPLICE *erb*B-2 protein or an analog thereof, i.e., the protein with an insertion, substitution or deletion mutation.

The invention further provides a method for preparing a SPLICE protein, or a related protein utilizing the purified and isolated nucleic acid molecules of the invention. In an embodiment a method for preparing a SPLICE *erb*B-2 protein is provided comprising (a) transferring a recombinant expression vector of the invention into a host cell; (b) selecting transformed host cells from untransformed host cells; (c) culturing a selected transformed host cell under conditions which allow expression of the SPLICE *erb*B-2 protein; and (d) isolating the SPLICE *erb*B-2 protein.

The invention further broadly contemplates a purified and isolated SPLICE *erb*B-2 protein which is capable of transforming Rat-1 fibroblasts. The SPLICE *erb*B-2 protein has an in frame deletion of sixteen (16) amino acids, two of which are conserved cysteine residues and an approximate molecular weight of 50KD. In an embodiment of the invention, a purified SPLICE *erb*B-2 protein is provided which has the amino acid sequence as shown in Figure 2 or SEQ ID NO:2. The invention also includes truncations of the protein and analogs, homologs, and isoforms of the protein and truncations thereof.

The SPLICE *erb*B-2 proteins of the invention may be conjugated with other molecules, such as proteins to prepare fusion proteins. This may be accomplished, for example, by the synthesis of N-terminal or C-terminal fusion proteins.

The invention further contemplates antibodies having specificity against an epitope of a SPLICE *erb*B-2 protein of the invention. Antibodies may be labelled with a detectable substance and they may be used to detect the SPLICE *erb*B-2 protein of the invention in tissues and cells.

The invention also permits the construction of nucleotide probes which are unique to the nucleic acid molecules of the invention and accordingly to a SPLICE *erb*B-2 protein of the invention. Thus, the invention also relates to a probe comprising a sequence encoding a SPLICE *erb*B-2 protein or fragment thereof. The probe may be labelled, for example, with a

detectable substance and it may be used to select from a mixture of nucleotide sequences a nucleotide sequence coding for a protein which displays one or more of the properties of the protein of the invention. The probes may also be used to detect a nucleic acid encoding a SPLICE *erb*B-2 protein of the invention in tissues and cells. In particular, the probes may be used as a diagnostic tool on biopsied tissues to assess the transformed state of the cell.

The SPLICE *erb*B-2 protein is likely involved in the induction of human breast tumors. Accordingly, in one embodiment the present invention provides a method of modulating SPLICE *erb*B-2 activity comprising administering an effective amount of inhibitor of SPLICE *erb*B-2 protein activity to a cell or animal in need thereof. In such an embodiment, it may be desirable to inhibit the expression of the SPLICE *erb*B-2 protein. Accordingly, the present invention provides a method of inhibiting or reducing cell transformation comprising administering an agent that inhibits the expression or activity of the SPLICE *erb*B-2 protein to an animal or cell in need thereof. Agents that may inhibit the SPLICE *erb*B-2 protein include antibodies to the protein or antisense oligonucleotides complimentary to a portion of a nucleic acid sequence encoding the protein.

The invention still further provides a method for identifying a substance which is capable of binding to a SPLICE *erb*B-2 protein or inactive form thereof, comprising reacting the SPLICE *erb*B-2 protein, or inactivated form, with at least one substance which potentially can bind with the SPLICE *erb*B-2 protein or an inactivated form thereof, under conditions which permit the formation of complexes between the substance and the protein or an inactivated form thereof, and assaying for complexes, for free substance, for non-complexed SPLICE *erb*B-2 pr inactivated form or for activation of SPLICE *erb*B-2 Specifically, a yeast two hybrid assay system may be utilized as a method for identifying proteins which interact with the protein (Fields, S. and Song, O. 1989, Nature, 34:245-247).

Still further, the invention provides a method for assaying a medium for the presence of an agonist or antagonist of the interaction of a SPLICE *erb*B-2 protein or an activated form thereof, and a substance which binds to the protein or an activated form thereof. In one embodiment, the method comprises providing a known concentration of a SPLICE *erb*B-2 protein, with a substance which is capable of binding to the protein and a suspected agonist or antagonist substance under conditions which permit the formation of complexes between the substance and SPLICE *erb*B-2 protein, and assaying for complexes, free substance, non-complexed SPLICE *erb*B-2 protein, or for activation of the protein.

Substances which affect the SPLICE *erb*B-2 protein may also be identified using the methods of the invention by comparing the pattern and level of expression of the protein in tissues and cells in the presence, and in the absence of the substance.

Antibodies to SPLICE *erb*B-2 protein or antisense oligonucleotides complimentary to a nucleic acid encoding the SPLICE *erb*B-2 protein as well as substances identified using a method of the invention may be used in the treatment of conditions involving cell

transformation such as cancer. Accordingly, the substances may be formulated into pharmaceutical compositions for adminstration to individuals suffering from cancer and preferably breast cancer. Accordingly, the substances may be formulated into pharmaceutical compositions for administration to individuals suffering from cancer.

Other features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the cDNA sequence of the SPLICE *erb*B-2 protein of the invention.

Figure 2 shows the amino acid sequence of SPLICE *erb*B-2 protein.

Figure 3a provides a schematic representation of a small region of the human *erb*B-2 genomic locus indicating the exon-intron boundaries.

Figure 3b shows an alignment of the wild-type and alternatively spliced *erb*B-2 messages indicating the sequences removed by the aberrant splicing event (grey box).

Figure 3c shows alternate mRNA splicing results in an *erb*B-2 transcript harboring an in-frame deletion.

Figure 4a shows partial amino acid sequence alignment of the wild-type and altered *erb*B-2 receptors tested in Rat-1 focus assays.

Figure 4b shows representative plates from focus assay #1 (see Table 1) illustrating the relative transforming abilities of the various forms of *erb*B-2 indicated in Figure 4a.

Figure 5a shows immunoblots of lysates from stable cell lines expressing wild-type *Erb*B-2 (WT) or the three altered receptors (VE, ECD. SPLICE).

Figure 5b shows immunoblots of lysates from stable cell lines expressing the various *Erb*B-2 receptors.

DETAILED DESCRIPTION OF THE INVENTION

The following standard abbreviations for amino acid residues are used throughout the specification: A, Ala - alanine; C, Cys - cysteine; D, Asp- aspartic acid; E, Glu - glutamic acid; F, Phe - phenylalanine; G, Gly - glycine; H, His - histidine; I, Ile - isoleucine; K, Lys - lysine; L, Leu - leucine; M, Met - methionine; N, Asn - asparagine; P, Pro - proline; Q, Gln - glutamine; R, Arg - arginine; S, Ser - serine; T, Thr - threonine; V, Val - valine; W, Trp-tryptophan; Y, Tyr - tyrosine; and p.Y., P.Tyr - phosphotyrosine.

I. Nucleic Acid Molecules Encoding SPLICE *erb*B-2

a. The Nucleic Acid Molecules

As mentioned above, the inventi n provides an isolated and purified nucleic acid molecule having a sequence encoding a SPLICE *erb*B-2 protein. The term "isolated and

purified" refers to a nucleic acid substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical precursors, or other chemicals when chemically synthesized. The term "nucleic acid" is intended to include DNA and RNA and can be either double stranded or single stranded.

The nucleic acid sequence of the cDNA encoding the SPLICE *erb*B-2 protein is shown in Figure 1 and SEQ.ID.NO.:1.

It will be appreciated that the invention includes nucleic acid molecules encoding truncations of SPLICE *erb*B-2 protein, and analogs and homologs of the protein and truncations thereof, as described herein. It will further be appreciated that variant forms of nucleic acid molecules of the invention which arise by alternative splicing of an mRNA corresponding to a cDNA of the invention are encompassed by the invention.

Another aspect of the invention provides a nucleic acid molecule which hybridizes under high stringency conditions to a nucleic acid molecule which comprises a sequence which encodes SPLICE *erb*B-2 protein having the amino acid sequence shown in Figure 2. Appropriate stringency conditions which promote DNA hybridization are known to those skilled in the art, or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C may be employed. The stringency may be selected based on the conditions used in the wash step. By way of example, the salt concentration in the wash step can be selected from a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be at high stringency conditions, at about 65°C.

Isolated and purified nucleic acid molecules encoding a protein having the activity of SPLICE *erb*B-2 protein as described herein, and having a sequence which differs from the nucleic acid sequence shown in Figure 1, due to degeneracy in the genetic code are also within the scope of the invention.

In addition, DNA sequence polymorphisms within the nucleotide sequence of SPLICE *erb*B-2 protein (especially those within the third base of a codon) may result in "silent" mutations in the DNA which do not affect the amino acid encoded. However, DNA sequence polymorphisms may lead to changes in the amino acid sequences of the protein within a population. It will be appreciated by one skilled in the art that these variations in one or more nucleotides of the nucleic acids encoding proteins having the activity of the SPLICE *erb*B-2 protein may exist among individuals within a population due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of the invention.

b. Isolation of Purified Nucleic Acid Molecules

An isolated and purified nucleic acid molecule of the invention which comprises DNA can be isolated by preparing a labelled nucleic acid probe based on all or part of the

nucleic acid sequence shown in Figure 1 and using this labelled nucleic acid probe to screen an appropriate DNA library (e.g. a cDNA or genomic DNA library). For instance, a cDNA library made from lymphocytes such as B cells can be used to isolate a cDNA encoding a protein having SPLICE *erb*B-2 activity by screening the library with the labelled probe using standard techniques. Alternatively, a genomic DNA library can be similarly screened to isolate a genomic clone encompassing a gene encoding a SPLICE *erb*B-2 protein. Nucleic acids isolated by screening of a cDNA or genomic DNA library can be sequenced by standard techniques.

An isolated and purified nucleic acid molecule of the invention which is DNA can also be isolated by selectively amplifying a nucleic acid encoding a SPLICE *erb*B-2 protein using the polymerase chain reaction (PCR) methods and cDNA or genomic DNA. It is possible to design synthetic oligonucleotide primers from the nucleotide sequence shown in Figure 1. A nucleic acid can be amplified from cDNA or genomic DNA using these oligonucleotide primers and standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. It will be appreciated that cDNA may be prepared from mRNA, by isolating total cellular mRNA by a variety of techniques, for example, by using the guanidinium-thiocyanate extraction procedure of Chirgwin et al., Biochemistry, 18, 5294-5299 (1979). cDNA is then synthesized from the mRNA using reverse transcriptase (for example, Moloney MLV reverse transcriptase available from Gibco/BRL, Bethesda, MD, or AMV reverse transcriptase available from Seikagaku America, Inc., St. Petersburg, FL).

An isolated and purified nucleic acid molecule of the invention which is RNA can be isolated by cloning a cDNA encoding a SPLICE *erb*B-2 protein into an appropriate vector which allows for transcription of the cDNA to produce an RNA molecule which encodes a protein which exhibits SPLICE *erb*B-2 protein activity. For example, a cDNA can be cloned downstream of a bacteriophage promoter, (e.g. a T7 promoter) in a vector, cDNA can be transcribed in vitro with T7 polymerase, and the resultant RNA can be isolated by standard techniques.

c. Chemical Synthesis of Nucleic Acid Molecules

A nucleic acid molecule of the invention may also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071).

Determination of whether a particular nucleic acid molecule encodes a protein having SPLICE *erb*B-2 protein activity can be accomplished by expressing the cDNA in an appropriate host cell by standard techniques, and testing the ability of the expressed protein

to transform Rat-1 fibroblasts. A cDNA having the biological activity of a SPLICE *erb*B-2 protein so isolated can be sequenced by standard techniques, such as dideoxynucleotide chain termination or Maxam-Gilbert chemical sequencing, to determine the nucleic acid sequence and the predicted amino acid sequence of the encoded protein.

d. Antisense Nucleic Acid Molecules

The sequence of a nucleic acid molecule of the invention may be inverted relative to its normal presentation for transcription to produce an antisense nucleic acid molecule. An antisense nucleic acid molecule may be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art.

II. SPLICE *erb*B-2 Proteins

As mentioned above, the present invention relates to an isolated SPLICE *erb*B-2 protein having an in frame deletion of sixteen (16) amino acids, two of which are conserved cysteine residues. The amino acid sequence of a SPLICE *erb*B-2 protein, Seq I.D. No. 2, of the present invention is shown in Figure 2.

The inventors have demonstrated that SPLICE *erb*B-2 protein is expressed in human breast tissue; in low levels in normal tissue and high levels in tumor tissue. The inventors have also demonstrated that the SPLICE *erb*B-2 protein is constitutively tyrosine phosphorated, and that this protein is capable of transforming Rat-1 fibroblasts.

The proteins of the present invention include truncations of the SPLICE *erb*B-2 protein, and analogs, and homologs and truncations thereof as described herein.

Truncated proteins may have an amino group (-NH2), a hydrophobic group (for example, carbobenzoxyl, dansyl, or T-butyloxycarbonyl), an acetyl group, a 9-fluorenylmethoxy-carbonyl (PMOC) group, or a macromolecule including but not limited to lipid-fatty acid conjugates, polyethylene glycol, or carbohydrates at the amino terminal end. The truncated proteins may have a carboxyl group, an amino group, a T-butyloxycarbonyl group, or a macromolecule including but not limited to lipid-fatty acid conjugates, polyethylene glycol, or carbohydrates at the carboxy terminal end.

Proteins of the invention may also include analogs of the SPLICE *erb*B-2 protein as shown in Figure 2 or truncations thereof as described herein, which may include, but are not limited to one or more amino acid substitutions, insertions, and/or deletions. Amino acid substitutions may be of a conserved or non-conserved nature. Conserved amino acid substitutions involve replacing one or more amino acids of the SPLICE *erb*B-2 amino acid sequence with amino acids of similar charge, size, and/or hydrophobicity characteristics. When only conserved substitutions are made the resulting analog should be functionally equivalent to the SPLICE *erb*B-2 protein. Non-conserved substitutions involve replacing one or more amino acids of the SPLICE *erb*B-2 amino acid sequence with one or more amino acids which possess dissimilar charge, size, and/or hydrophobicity characteristics.

One or more amino acid insertions may be introduced into the SPLICE *erb*B-2 protein. Amino acid insertions may consist of single amino acid residues or sequential amino acids ranging from 2 to 15 amino acids in length.

The proteins of the invention also include homologs of the SPLICE *erb*B-2 protein (Figure 2) and/or truncations thereof as described herein. Such homologs are proteins whose amino acid sequences are comprised of the amino acid sequences of SPLICE *erb*B-2 protein regions from other species that hybridize under stringent hybridization conditions with a probe used to obtain the protein.

The invention also contemplates isoforms of the protein of the invention. An isoform contains the same number and kinds of amino acids as the protein of the invention or subtle spliced differences which alter the overall number of amino acids resulting in a slight shift in molecular weight. The isoforms contemplated by the present invention are those having the same properties as a SPLICE *erb*B-2 protein of the invention as described herein.

The present invention also includes SPLICE *erb*B-2 protein conjugated with a selected protein, or a selectable marker protein (see below) to produce fusion proteins. Further, the present invention also includes activated or phosphorylated SPLICE *erb*B-2 proteins of the invention. Additionally, immunogenic portions of SPLICE *erb*B-2 proteins are within the scope of the invention.

Proteins of the invention may be prepared using recombinant DNA methods. Accordingly, the nucleic acid molecules of the present invention having a sequence which encodes a SPLICE *erb*B-2 protein of the invention may be incorporated in a known manner into an appropriate expression vector which ensures good expression of the protein. Possible expression vectors include but are not limited to cosmids, plasmids, or modified viruses (e.g. replication defective retroviruses, adenoviruses and adeno-associated viruses), so long as the vector is compatible with the host cell used. The expression vectors are "suitable for transformation of a host cell", means that the expression vectors contain a nucleic acid molecule of the invention and regulatory sequences selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid molecule. Operatively linked is intended to mean that the nucleic acid is linked to regulatory sequences in a manner which allows expression of the nucleic acid.

The invention therefore contemplates a recombinant expression vector of the invention containing a nucleic acid molecule of the invention, or a fragment thereof, and the necessary regulatory sequences for the transcription and translation of the inserted protein-sequence. Suitable regulatory sequences may be derived from a variety of sources, including mammalian, bacterial, fungal, viral or insect genes (For example, see the regulatory sequences described in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Selection of appropriate regulatory sequences is dependent on the host cell chosen as discussed below, and may be readily

accomplished by one of ordinary skill in the art. Examples of such regulatory sequences include: a transcriptional promoter and enhancer or RNA polymerase binding sequence, a ribosomal binding sequence, including a translation initiation signal. Additionally, depending on the host cell chosen and the vector employed, other sequences, such as an origin of replication, additional DNA restriction sites, enhancers, and sequences conferring inducibility of transcription may be incorporated into the expression vector. It will also be appreciated that the necessary regulatory sequences may be supplied by the native the SPLICE *erb*B-2 gene and/or its flanking regions.

The invention further provides a recombinant expression vector comprising a DNA nucleic acid molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression, by transcription of the DNA molecule, of an RNA molecule which is antisense to the nucleotide sequence of SEQ ID NO: 1. Regulatory sequences operatively linked to the antisense nucleic acid can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance a viral promoter and/or enhancer, or regulatory sequences can be chosen which direct tissue or cell type specific expression of antisense RNA. Antisense oligodeoxynucleotides seem promising for the purposes of inhibiting expression of genes as long as they are efficiently protected against degradation and targetted into cells. Supramolecular biovectors (SMBV) are antisense oligonucleotide carriers which improve both cellular uptake and protection of oligodeoxynucleotide (see for example Allal, C. et al. Br,. J. Cancer 77:1448-1453 (1998)) and are further provided as part of the present invention.

The recombinant expression vectors of the invention may also contain a selectable marker gene which facilitates the selection of host cells transformed or transfected with a recombinant molecule of the invention. Examples of selectable marker genes are genes encoding a protein such as G418 and hygromycin which confer resistance to certain drugs, β-galactosidase, chloramphenicol acetyltransferase, firefly luciferase, green fluorescent protein, alkaline phosphatase, yellow fluorescent protein, blue fluorescent protein or an immunoglobulin or portion thereof such as the Fc portion of an immunoglobulin preferably IgG. Transcription of the selectable marker gene is monitored by changes in the concentration of the selectable marker protein such as β-galactosidase, chloramphenicol acetyltransferase, or firefly luciferase. If the selectable marker gene encodes a protein conferring antibiotic resistance such as neomycin resistance, transformant cells can be selected with G418. Cells that have incorporated the selectable marker gene will survive, while the other cells die. This makes it possible to visualize and assay for expression of recombinant expression vectors of the invention and in particular to determine the effect of a mutation on expression and phenotype. It will be appreciated that selectable markers can be introduced on a separate vector from the nucleic acid of interest.

The recombinant expression vectors may also contain genes which encode a fusion moiety which provides increased expression of the recombinant protein; increased solubility of the recombinant protein; and aid in the purification of the target recombinant protein by acting as a ligand in affinity purification. For example, a proteolytic cleavage site may be added to the target recombinant protein to allow separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Typical fusion expression vectors include pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the recombinant protein.

Recombinant expression vectors can be introduced into host cells to produce a transformant host cell. The term "transformant host cell" is intended to include prokaryotic and eukaryotic cells which have been transformed or transfected with a recombinant expression vector of the invention. The terms "transformed with", "transfected with", "transformation" and "transfection" are intended to encompass introduction of nucleic acid (e.g. a vector) into a cell by one of many possible techniques known in the art. Prokaryotic cells can be transformed with nucleic acid by, for example, electroporation or calcium-chloride mediated transformation. Nucleic acid can be introduced into mammalian cells via conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofectin, electroporation or microinjection. Suitable methods for transforming and transfecting host cells can be found in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

Suitable host cells include a wide variety of prokaryotic and eukaryotic host cells. For example, the PKT proteins of the invention may be expressed in mammalian cells, bacterial cells such as *E. coli*, insect cells (using baculovirus) or yeast cells. Other suitable host cells can be found in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1991).

Mammalian cells suitable for carrying out the present invention include, among others: Cos-7 (green monkey kidney), Saos-2 (human osteosarcoma), PC12, NIH-3T3, MONC, SY5Y, P19, G361 (human melanoma), A549 (human lung carcinoma), SW480 (human colorectal adenocarcinoma), Raji (human Burkitt's lymphoma), MOLT-4 (human lymphoblastic leukemia), K562 (human chronic myelogenous leukemia), S3 (Hela cells), HL-60 (human promyelocytic leukemia, and breast cancer cell lines; T47Ds, MCF-7s and C127s, COS (e.g., ATCC No. CRL 1650 or 1651), BHK (e.g., ATCC No. CRL 6281), CHO (ATCC No. CCL 61), HeLa (e.g., ATCC No. CCL 2), 293 (ATCC No. 1573) and NS-1 cells. Suitable expression vectors for directing expression in mammalian cells generally include a promoter (e.g., derived from viral material such as polyoma, Adenovirus 2, cytomegalovirus, retrovirus (pBabe and LXHSD), and Simian Virus 40), as well as other transcriptional and

translational control sequences. Examples of mammalian expression vectors include pCDM8 (Seed, B., (1987) Nature 329:840) and pMT2PC (Kaufman et al. (1987), EMBOJ. 6:187-195).

More particularly, bacterial host cells suitable for carrying out the present invention include *E. coli, B. subtilis, Salmonella typhimurium*, and various species within the genus' *Pseudomonas, Streptomyces*, and *Staphylococcus*, as well as many other bacterial species well known to one of ordinary skill in the art. Suitable bacterial expression vectors preferably comprise a promoter which functions in the host cell, one or more selectable phenotypic markers, and a bacterial origin of replication. Representative promoters include the β-lactamase (penicillinase) and lactose promoter system (see Chang et al., Nature 275:615, 1978), the trp promoter (Nichols and Yanofsky, Meth in Enzymology 101:155, 1983) and the tac promoter (Russell et al., Gene 20: 231, 1982). Representative selectable markers include various antibiotic resistance markers such as the kanamycin or ampicillin resistance genes. Suitable expression vectors include but are not limited to bacteriophages such as lambda derivatives or plasmids such as pBR322 (see Bolivar et al., Gene 2:9S, 1977), the pUC plasmids pUC18, pUC19, pUC118, pUC119 (see Messing, Meth in Enzymology 101:20-77, 1983 and Vieira and Messing, Gene 19:259-268, 1982), and pNH8A, pNH16a, pNH18a, and Bluescript M13 (Stratagene, La Jolla, Calif.). Typical fusion expression vectors which may be used are discussed above, e.g. pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ). Examples of inducible non-fusion expression vectors include pTrc (Amann et al., (1988) Gene 69:301-315) and pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 60-89).

Yeast and fungi host cells suitable for carrying out the present invention include, but are not limited to *Saccharomyces cerevisae*, the genera *Pichia* or *Kluyveromyces* and various species of the genus *Aspergillus*. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari. et al., (1987) Embo J. 6:229-234), pMFa (Kurjan and Herskowitz, (1982) Cell 30:933-943), pJRY88 (Schultz et al., (1987) Gene 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Protocols for the transformation of yeast and fungi are well known to those of ordinary skill in the art (see Hinnen et al., PNAS USA 75:1929, 1978; Itoh et al., J. Bacteriology 153:163, 1983, and Cullen et al. Bio/Technology 5:369, 1987).

Given the teachings provided herein, promoters, terminators, and methods for introducing expression vectors of an appropriate type into plant, avian, and insect cells may also be readily accomplished. For example, within one embodiment, the proteins of the invention may be expressed from plant cells (see Sinkar et al., J. Biosci (Bangalore) 11:47-58, 1987, which reviews the use of *Agrobacterium rhizogenes* vectors; see also Zambryski et al., Genetic Engineering, Principles and Methods, Hollaender and Setlow (eds.), Vol. VI, pp. 253-278, Plenum Press, New York, 1984, which describes the use of expression vectors for plant cells, including, among others, pAS2022, pAS2023, and pAS2034).

Insect cells suitable for carrying out the present invention include cells and cell lines from *Bombyx* or *Spodotera* species. Baculovirus vectors available for expression of proteins in cultured insect cells (SF 9, SF 21, and T.ni-High Five cells) include the pAc series (Smith et al., (1983) Mol. Cell Biol. 3:2156-2165) and the pVL series (Luckow, V.A., and Summers, M.D., (1989) Virology 170:31-39) and pBAC PAK.

Alternatively, the proteins of the invention may also be expressed in non-human transgenic animals such as mice, rats, rabbits, sheep and pigs (see Hammer et al. (Nature 315:680-683, 1985), Palmiter et al. (Science 222:809-814, 1983), Brinster et al. (Proc Natl. Acad. Sci USA 82:44384442, 1985), Palmiter and Brinster (Cell. 41:343-345, 1985) and U.S. Patent No. 4,736,866).

The invention further provides a recombinant expression vector for the transcription and translation in invertebrate animals including, but not limited to, zebrafish, xenopus and drosophila.

The proteins of the invention may also be prepared by chemical synthesis using techniques well known in the chemistry of proteins such as solid phase synthesis (Merrifield, 1964, J. Am. Chem. Assoc. 85:2149-2154) or synthesis in homogenous solution (Houbenweyl, 1987, Methods of Organic Chemistry, ed. E. Wansch, Vol. 15 I and II, Thieme, Stuttgart).

N-terminal or C-terminal fusion proteins comprising the SPLICE *erb*B-2 protein of the invention conjugated with other molecules, such as proteins may be prepared by fusing, through recombinant techniques, the N-terminal or C-terminal of the protein, and the sequence of a selected protein or selectable marker protein with a desired biological function. The resultant fusion proteins contain the SPLICE *erb*B-2 protein fused to the selected protein or marker protein as described herein. Examples of proteins which may be used to prepare fusion proteins include green fluorescent protein (GFP), yellow fluorescent protein, blue fluorescent protein, alkaline phosphatase, immunoglobulins, glutathione-S-transferase (GST), haemagglutonin (HA), and truncated myc.

Phosphorylation or activation SPLICE *erb*B-2 proteins of the invention may be ensured using, for example, the method described in Reedijk et al. The EMBO Journal 11(4):1365, 1992.

III. Utility of the Nucleic Acid Molecules and Proteins of the Invention

Diagnostic Uses

The nucleic acid molecules of the invention allow those skilled in the art to construct nucleotide probes for use in the detection of nucleic acid sequences in biological materials. Suitable probes include nucleic acid molecules based on nucleic acid sequences encoding at least 6 sequential amino acids from regions of the SPLICE *erb*B-2 protein as shown in Figure 2. A nucleotide probe may be labelled with a detectable substance such as a radioactive label which provides for an adequate signal and has sufficient half-life such as ^{32}P, ^{3}H, ^{14}C or the like. Other detectable substances which may be used include antigens that are recognized by

a specific labelled antibody, fluorescent compounds, enzymes, antibodies specific for a labelled antigen, and luminescent compounds. An appropriate label may be selected having regard to the rate of hybridization and binding of the probe to the nucleotide to be detected and the amount of nucleotide available for hybridization. Labelled probes may be hybridized to nucleic acids on solid supports such as nitrocellulose filters or nylon membranes as generally described in Sambrook et al, 1989, Molecular Cloning, A Laboratory Manual (2nd ed.). The nucleic acid probes may be used to detect genes, preferably in human cells, that encode the SPLICE *erb*B-2 protein. The nucleotide probes may be useful in the diagnosis of disorders of cell transformation such as cancer. The nucleotide probes may also be used as a diagnostic tool on tissue biopsies to assess the transformed state of the cell. The probes may also be used in *in situ* hybridization of early embryos to assess both the onset and pattern of expression during development.

SPLICE *erb*B-2 proteins of the invention can be used to prepare antibodies specific for the SPLICE *erb*B-2 proteins that may be used to detect the protein in a biological sample. Antibodies can be prepared which bind a distinct epitope in an unconserved region of the protein. An unconserved region of the protein is one which does not have substantial sequence homology to other proteins. Antibodies which distinguish between SPLICE *erb*B-2 and other *erb*B-2 proteins can be prepared which bind distinct epitopes of the protein.

Conventional methods can be used to prepare the antibodies. For example, by using a peptide of the SPLICE *erb*B-2 protein, polyclonal antisera or monoclonal antibodies can be made using standard methods. A mammal, (e.g., a mouse, hamster, or rabbit) can be immunized with an immunogenic form of the peptide which elicits an antibody response in the mammal. Techniques for conferring immunogenicity on a peptide include conjugation to carriers or other techniques well known in the art. For example, the protein or peptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassay procedures can be used with the immunogen as antigen to assess the levels of antibodies. Following immunization, antisera can be obtained and, if desired, polyclonal antibodies isolated from the sera.

To produce monoclonal antibodies, antibody producing cells (lymphocytes) can be harvested from an immunized animal and fused with myeloma cells by standard somatic cell fusion procedures thus immortalizing these cells and yielding hybridoma cells. Such techniques are well known in the art, (e.g., the hybridoma technique originally developed by Kohler and Milstein (Nature 256, 495-497 (1975)) as well as other techniques such as the human B-cell hybridoma technique (Kozbor et al., Immunol. Today 4, 72 (1983)), the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al. Monoclonal Antibodies in Cancer Therapy (1985) Allen R. Bliss, Inc., pages 77-96), and screening of combinatorial antibody libraries (Huse et al., Science 246, 1275 (1989)). Hybridoma cells can

be screened immunochemically for production of antibodies specifically reactive with the peptide and the monoclonal antibodies can be isolated. Therefore, the invention also contemplates hybridoma cells secreting monoclonal antibodies with specificity for the SPLICE *erb*B-2 protein as described herein.

The term "antibody" as used herein is intended to include fragments thereof which also specifically react with a SPLICE *erb*B-2 protein, or peptide thereof, having the activity of the SPLICE *erb*B-2 protein. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above. For example, F(ab')2 fragments can be generated by treating antibody with pepsin. The resulting F(ab')2 fragment can be treated to reduce disulfide bridges to produce Fab' fragments.

Chimeric antibody derivatives, i.e., antibody molecules that combine a non-human animal variable region and a human constant region are also contemplated within the scope of the invention. Chimeric antibody molecules can include, for example, the antigen binding domain from an antibody of a mouse, rat, or other species, with human constant regions. Conventional methods may be used to make chimeric antibodies containing the immunoglobulin variable region which recognizes the gene product of SPLICE *erb*B-2 antigens of the invention (See, for example, Morrison et al., Proc. Natl Acad. Sci. U.S.A. 81,6851 (1985); Takeda et al., Nature 314, 452 (1985), Cabilly et al., U.S. Patent No. 4,816,567; Boss et al., U.S. Patent No. 4,816,397; Tanaguchi et al., European Patent Publication EP171496; European Patent Publication 0173494, United Kingdom patent GB 2177096B). It is expected that chimeric antibodies would be less immunogenic in a human subject than the corresponding non-chimeric antibody.

Monoclonal or chimeric antibodies specifically reactive with a protein of the invention as described herein can be further humanized by producing human constant region chimeras, in which parts of the variable regions, particularly the conserved framework regions of the antigen-binding domain, are of human origin and only the hypervariable regions are of non-human origin. Such immunoglobulin molecules may be made by techniques known in the art, (e.g., Teng et al., Proc. Natl. Acad. Sci. U.S.A., 80, 7308-7312 (1983); Kozbor et al., Immunology Today, 4, 7279 (1983); Olsson et al., Meth. Enzymol., 92, 3-16 (1982)), and PCT Publication WO92/06193 or EP 0239400). Humanized antibodies can also be commercially produced (Scotgen Limited, 2 Holly Road, Twickenham, Middlesex, Great Britain.)

Specific antibodies, or antibody fragments, reactive against proteins of the invention may also be generated by screening expression libraries encoding immunoglobulin genes, or portions thereof, expressed in bacteria with peptides produced from the nucleic acid molecules of the present invention. For example, complete Fab fragments, VH regions and FV regions can be expressed in bacteria using phage expression libraries (See for example Ward et al., Nature 341, 544-546: (1989); Huse et al., Science 246, 1275-1281 (1989); and McCafferty

et al. Nature 348, 552-554 (1990)). Alternatively, a SCID-hu mouse, for example the model developed by Genpharm, can be used to produce antibodies, or fragments thereof.

Antibodies specifically reactive with the SPLICE *erb*B-2 protein, or derivatives thereof, such as enzyme conjugates or labeled derivatives, may be used to detect SPLICE *erb*B-2 protein in various biological materials, for example they may be used in any known immunoassays which rely on the binding interaction between an antigenic determinant of the protein, and the antibodies. Examples of such assays are radioimmunoassays, enzyme immunoassays (e.g. ELISA), immunofluorescence, immunoprecipitation, latex agglutination, hemagglutination, and histochemical tests. Thus, the antibodies may be used to detect and quantify the SPLICE *erb*B-2 protein in a sample in order to determine its role in particular cellular events or pathological states, and to diagnose and treat such pathological states. The antibodies may also be used on biopsied tissue to assess whether or not a cell is transformed.

In particular, antibodies of the invention may be used in immuno-histochemical analyses, for example, at the cellular and sub-subcellular level, to detect the SPLICE *erb*B-2 protein, to localise it to particular cells and tissues and to specific subcellular locations, and to quantitate the level of expression.

Cytochemical techniques known in the art for localizing antigens using light and electron microscopy may be used to detect the SPLICE *erb*B-2 protein. Generally, an antibody of the invention may be labelled with a detectable substance and the SPLICE *erb*B-2 protein may be localised in tissue based upon the presence of the detectable substance. Examples of detectable substances include various enzymes, fluorescent materials, luminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, biotin, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; and examples of suitable radioactive material include radioactive iodine I^{125}, I^{131} or tritium. Antibodies may also be coupled to electron dense substances, such as ferritin or colloidal gold, which are readily visualised by electron microscopy.

Indirect methods may also be employed in which the primary antigen-antibody reaction is amplified by the introduction of a second antibody, having specificity for the antibody reactive against a SPLICE *erb*B-2 protein. By way of example, if the antibody having specificity against the protein is a rabbit IgG antibody, the second antibody may be goat anti-rabbit gamma-globulin labelled with a detectable substance as described herein.

Where a radioactive label is used as a detectable substance, SPLICE *erb*B-2 protein may be localized by autoradiography. The results of autoradiography may be quantitated

by determining the density of particles in the autoradiographs by various optical methods, or by counting the grains.

As discussed herein,SPLICE *erb*B-2 protein is an *erb*B-2 receptor with intracellular tyrosine kinese activity which is constitutively tyrosine phosphorylated having an in frame deletion of 16 amino acids, two of which are conserved cysteine residues. This novel receptor is capable of transforming Rat-1 fibroblasts and as such likely plays a role in regulating transformation of cells. Therefore, the above described methods for detecting nucleic acid molecules and SPLICE *erb*B-2 proteins of the invention, can be used to monitor cell transformation. It would also be apparent to one skilled in the art that the above described methods may be used to study the developmental expression of the SPLICE *erb*B-2 protein and, accordingly, will provide further insight into the role of the SPLICE *erb*B-2 protein in cell development. In addition, since the SPLICE *erb*B-2 protein is highly expressed in mammary tumor cells, it may be involved in mammary tissue development or physiology.

Therapeutic Uses

The SPLICE *erb*B-2 protein of the invention is likely involved in the regulation of cell signalling pathways that control cell transformation. Accordingly, the present invention provides a method of modulating cell transformation comprising administering an effective amount of a SPLICE *erb*B-2 protein or a nucleic acid encoding a SPLICE *erb*B-2 protein to a cell or animal in need thereof. The term "effective amount" as used herein means an amount effective, at dosages and for periods of time necessary to achieve the desired results.

In another embodiment, the present invention provides a method of modulating tumor cell proliferation. In a preferred embodiment, the invention provides a method of inhibiting or reducing tumor cell proliferation by administering an agent that inhibits the expression or the biological activity of the SPLICE *erb*B-2 protein. Agents that inhibit the activity of a SPLICE *erb*B-2 protein include antibodies to the protein. Agents that inhibit the expression of the SPLICE *erb*B-2 gene include antisense oligonucleotides to a SPLICE *erb*B-2 nucleic acid sequence. Both of these are described above.

In addition to antibodies and antisense oligonucleotides, other substances that inhibit SPLICE *erb*B-2 protein expression or activity may also be identified.

Substances which affect SPLICE *erb*B-2 protein activity can be identified based on their ability to bind to the SPLICE *erb*B-2 protein. Therefore, the invention also provides methods for identifying substances which are capable of binding to a SPLICE *erb*B-2 protein. In particular, such methods may be used to identify substances capable of binding to, and in some cases activating (i.e., phosphorylating), or deactivating the protein of the invention.

Substances which can bind with a SPLICE *erb*B-2 protein of the invention may be identified by reacting the protein with a substance which potentially binds to the protein, and assaying for complexes, for free substance, or for non-complexed SPLICE *erb*B-2 protein, or for activation or deactivation (as the case may be) of the protein. In particular, a yeast two

hybrid assay system may be used to identify proteins which interact with the SPLICE *ERBB*-2 protein (Fields, S. and Song, O., 1989, Nature, 340:245-247).

Conditions which permit the formation of a substance and SPLICE *erb*B-2 protein complexes may be selected having regard to factors such as the nature and amounts of the substance and the protein.

The substance-protein complex, free substance or non-complexed proteins may be isolated by conventional isolation techniques, for example, salting out, chromatography, electrophoresis, gel filtration, fractionation, absorption, polyacrylamide gel electrophoresis, agglutination, or combinations thereof. To facilitate the assay of the components, antibody against the SPLICE *erb*B-2 protein or the substance, or labelled SPLICE *erb*B-2 protein, or a labelled substance may be utilized. The antibodies, proteins, or substances may be labelled with a detectable substance as described above.

Substances which bind to and activate or deactivate SPLICE *erb*B-2 protein of the invention may be identified by assaying for phosphorylation of the tyrosine residues of the protein.

SPLICE *erb*B-2 protein, or the substance used in the method of the invention may be immobilized. For example, the SPLICE *erb*B-2 protein or substance may be bound to a suitable carrier. Examples of suitable carriers are agarose, cellulose, dextran, Sephadex, Sepharose, carboxymethyl cellulose polystyrene, filter paper, ion-exchange resin, plastic film, plastic tube, glass beads, polyamine-methyl vinyl-ether-maleic acid copolymer, amino acid copolymer, ethylene-maleic acid copolymer, nylon, silk, etc. The carrier may be in the shape of, for example, a tube, test plate, beads, disc, sphere etc.

The immobilized protein or substance may be prepared by reacting the material with a suitable insoluble carrier using known chemical or physical methods, for example, cyanogen bromide coupling.

The proteins or substance may also be expressed on the surface of a cell using the methods described herein.

The invention also contemplates a method for assaying for an agonist or antagonist of the binding of the SPLICE *erb*B-2 protein with a substance which is capable of binding with the SPLICE *erb*B-2 protein. The agonist or antagonist may be an endogenous physiological substance or it may be a natural or synthetic substance. Substances which are capable of binding with the SPLICE *erb*B-2 protein may be identified using the methods set forth herein.

It will be understood that the agonists and antagonists that can be assayed using the methods of the invention may act on one or more of the binding sites on the protein or substance including agonist binding sites, competitive antagonist binding sites, non-competitive antagonist binding sites or allosteric sites.

The invention also makes it possible to screen for antagonists that inhibit the effects of an agonist of the interaction of the SPLICE *erb*B-2 protein with a substance which is capable of binding to the protein. Thus, the invention may be used to assay for a substance that competes for the same binding site of the SPLICE *erb*B-2 protein.

The methods described above may be used to identify a substance which is capable of binding to an activated SPLICE *erb*B-2 protein, and to assay for an agonist or antagonist of the binding of activated protein, with a substance which is capable of binding with activated SPLICE *erb*B-2 protein. An activated (i.e., phosphorylated) SPLICE *erb*B-2 protein may be prepared using the methods described (for example in Reedijk et al. The EMBO Journal, 11(4):1365, 1992) for producing a constitutively tyrosine phosphorylated protein.

It will also be appreciated that intracellular substances which are capable of binding to SPLICE *erb*B-2 protein may be identified using the methods described herein.

The invention further provides a method for assaying for a substance that affects an SPLICE *erb*B-2 protein regulatory pathway comprising administering to a non-human animal or to a tissue of an animal, a substance suspected of affecting a SPLICE *erb*B-2 protein regulatory pathway, and quantitating the protein or nucleic acids encoding the protein, or examining the pattern and/or level of expression of SPLICE *erb*B-2 protein, in the non-human animal or tissue. SPLICE *erb*B-2 protein may be quantitated and its expression may be examined using the methods described herein.

Substances identified by the methods described herein, may be used for modulating SPLICE *erb*B-2 protein regulatory pathways and accordingly may be used in the treatment of conditions involving perturbation of SPLICE *erb*B-2 protein signalling pathways. In particular, the substances may be particularly useful in the treatment of disorders of cell proliferation such as cancer.

As stated previously, SPLICE *erb*B-2 protein is involved in cell transformation and inhibitors of this protein may be useful in reversing disorders involving cell transformation such as cancer. In contrast, stimulators of the SPLICE *erb*B-2 protein may be useful in the treatment of disorders requiring stimulation of proliferation. Accordingly, substances that stimulate SPLICE *erb*B-2 protein (for example, identified using the methods of the invention) may be used to stimulate cell proliferation. Substances which stimulate cell proliferation may be useful in the treatment of conditions involving damaged cells including conditions in which degeneration of tissue occurs such as arthropathy, bone resorption, inflammatory disease, degenerative disorders of the central nervous system and for promoting wound healing.

<u>Pharmaceutical Compositions</u>

All of the above described substances (such as the SPLICE *erb*B-2 protein, the nucleic acid encoding such protein, antibodies to such protein, antisense oligonucleotides to the

nucleic acid molecules and substances that m dulate SPLICE *erb*B-2 protein activity) may be formulated into pharmaceutical compositions for adminstration to subjects in a biologically compatible form suitable for administration *in vivo*. By "biologically compatible form suitable for administration *in vivo*" is meant a form of the substance to be administered in which any toxic effects are outweighed by the therapeutic effects. The substances may be administered to living organisms including humans, and animals. In this context and throughout this specification animal means any member of the animal kingdom. Administration of a therapeutically active amount of pharmaceutical compositions of the present invention is defined as an amount effective, at dosages and for periods of time necessary to achieve the desired result. For example, a therapeutically active amount of a substance may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of antibody to elicit a desired response in the individual. Dosage regima may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation.

Active substance may be administered in a convenient manner such as by injection (subcutaneous, intravenous, etc.), oral administration, inhalation, transdermal application, or rectal administration. Depending on the route of administration, the active substance may be coated in a material to protect the compound from the action of enzymes, acids and other natural conditions which may inactivate the compound. If the active substance is a nucleic acid encoding a SPLICE *erb*B-2 protein or an antisense oligonucleotide it may be delivered using techniques known in the art. Recombinant molecules comprising an antisense sequence or oligonucleotide may be directly introduced into cells or tissues *in vivo* using delivery vehicles such as retroviral vectors, adeno viral vectors and DNA virus vectors. They may also be introduced using physical techniques such as microinjection and electroporation or chemical methods such as co-precipitation and incorporation of DNA into liposomes.

The compositions described herein can be prepared by per se known methods for the preparation of pharmaceutically acceptable compositions which can be administered to subjects, such that an effective quantity of the active substance is combined in a mixture with a pharmaceutically acceptable vehicle. Suitable vehicles are described, for example, in Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA 1985). On this basis, the compositions include, albeit not exclusively, solutions of the substances in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered solutions with a suitable pH and iso-osmotic with the physiological fluids.

Reagents suitable for applying methods of the invention to identify substances that affect SPLICE *erb*B-2 protein may be packaged into convenient kits providing the necessary

materials packaged into suitable containers. The kits may also include suitable supports useful in performing the methods of the invention.

Experimental Models

The invention also provides methods for studying the function of SPLICE *erb*B-2 protein. Cells, tissues and non-human animals that express or over-express SPLICE *erb*B-2 protein may be prepared by transfecting cells, tissues or oocytes (to prepare transgenic animals) with a recombinant expression vector of the invention (as described previously). In particular, transgenic technology (via nuclear oocyte microinjection of naked DNA) will assay the effect of over expression or alterations of the SPLICE *erb*B-2 protein expression in various developmental systems, including bone development, neurogenesis, mammary development, lung epithelial development.

Cells, tissues, and non-human animals lacking in SPLICE *erb*B-2 protein expression or partially lacking in SPLICE *erb*B-2 protein expression may be developed using recombinant expression vectors of the invention having specific deletion or insertion mutations in the SPLICE *erb*B-2 gene. A recombinant expression vector may be used to inactivate or alter the endogenous gene by homologous recombination, and thereby create a SPLICE *erb*B-2 protein deficient cell, tissue or animal.

Null alleles may be generated in cells, such as embryonic stem cells by deletion mutation. A recombinant SPLICE *erb*B-2 gene may also be engineered to contain an insertion mutation which inactivates the SPLICE *erb*B-2 protein. Such a construct may then be introduced into a cell, such as an embryonic stem cell, by a technique such as transfection, electroporation, injection etc. Cells lacking an intact SPLICE *erb*B-2 gene may then be identified, for example by Southern blotting, Northern Blotting or by assaying for SPLICE *erb*B-2 protein using the methods described herein. Such cells may then be fused to embryonic stem cells to generate transgenic non-human animals deficient in SPLICE *erb*B-2 protein. Germline transmission of the mutation may be achieved, for example, by aggregating the embryonic stem cells with early stage embryos, such as 8 cell embryos, *in vitro*; transferring the resulting blastocysts into recipient females and; generating germline transmission of the resulting aggregation chimeras. Such a mutant animal may be used to define specific cell populations, developmental patterns and *in vivo* processes, normally dependent on SPLICE *erb*B-2 protein expression.

The present invention also includes the preparation of tissue specific knock-outs of the SPLICE *erb*B-2 gene.

EXAMPLES

EXAMPLE 1

Cloning and Sequencing of SPLICE *erb*B-2 Gene/Protein

DNA constructions

Expression plasmids containing various human *erbB-2* cDNAs were constructed by inserting each cDNA as a *Hin*d III fragment into the corresponding site of pJ4Ω. The cDNAs encoding both wild-type and activated *erbB-2* (valine to glutamic acid substitution in the transmembrane domain) were generous gifts of Nancy E. Hynes (Friedrich Miescher Institute, Basel, Switzerland). The *erbB-2* (ECD DEL) mutant was generated by oligonucleotide-directed mutagenesis and the *erbB-2* (SPLICE) mutation was amplified from human breast tumor RNA. The polymerase chain reaction (PCR) products were sequenced to ensure that only the desired mutations were present.

The sequence of the cDNA is shown in Figure 1 and SEQ.ID.NO. 1. The amino acid sequence is shown in Figure 2 and SEQ.ID.NO. 2.

EXAMPLE 2

RT-PCR of total RNA from Normal and Tumor Breast Tissue

Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

The reverse transcription reaction was performed using total RNA isolated from a human prostate carcinoma cell line (DU145) as previously described (Siegel *et al.*, 1994). After incubation at 37°C for 2 h, 1.0 μl of the RT mix was added to the following: 0.5 μl of each oligonucleotide primer (10μM), 2.0 μl of 10X PCR buffer (200 mM Tris-HCl [pH 8.4], 500 mM KCl), 1.0 μl of 50 mM $MgCl_2$, 1.0 μl of deoxynucleotide triphosphate (dNTP) mix (10mM [each] dATP, dCTP, dGTP, and dTTP), 1.0 μl of [a-^{32}P]dCTP (10 μCi) and distilled water to bring the final volume to 20.0 μl. The reaction mixture was overlaid with 10.0 μl of paraffin oil and heated to 95°C for 5 min. The temperature was dropped to 80°C for 2 min and 5.0 μl of the following mixture was added to each sample: 0.1 μl of *Taq* DNA polymerase (GIBCO BRL), 0.5 μl of 10X PCR buffer, and 4.4 μl of distilled water. PCR amplification was performed for 30 cycles of 1 min at 94°C, 30 s at 55°C, and 1 min at 72°C. Both the wild-type and alternately spliced form of *erbB-2* wefe amplified with the following primers; AB12904 (TTTCCAGATGAGGAGGGC) (SEQ. I.D. NO. 3) and AB12903 (CGGAATTCCTGTCACCAGCTGCACCGT) (SEQ. I.D. No. 4), which amplify the region corresponding to nucleotides 1996-2544 of the human *erbB-2* cDNA (Coussens *et al.*,1985).

Inspection of the intron/exon structure of both the murine and human *erb*B-2 genomic loci revealed that an in-frame deletion could result from the removal of an exon immediately preceding the transmembrane domain (see Figure 3*a*). This alternatively spliced product would contain a deletion that closely resembles the transgene deletions observed in the MMTW wild-type *neu* transgenic mice (Figure 3*b*). To assess whether this putative splice form could be detected in human breast tumors, total RNA derived from both normal human

breast tissue and tumor samples were subjected to reverse transcriptase/polymerase chain reaction (RT-PCR) with oligonucleotide primers flanking the region of interest. The results of these analyses revealed that both normal and tumor tissue displayed evidence of the spliced transcript (Figure 3c). Direct sequence analyses of selected RT-PCR products confirmed that this altered spliced form possessed the expected 16 amino acid in-frame deletion (Figure 3*b*).

EXAMPLE 3

Transformation of Rat-1 Fibroblasts

Cell culture and Focus assays

Rat-1 fibroblasts were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), penicillin, streptomycin, and fungizone (GIBCO BRL). Stable cell lines were derived by electroporation as described previously (Siegel *et al.*, 1994). DNA was introduced into the Rat-1 cells at a 50:1 ratio of expression plasmid to puromycin resistance plasmid (PGK-puro) and resistant colonies were selected for 10 days in media containing puromycin (3.0 μg/ml) prior to deriving clonal cell lines.

The *erbB*-2 focus assays were performed essentially as described (Siegel *et al.*, 1994) with the exception that monolayers were maintained for 12 days in DMEM supplemented with 2% FBS, penicillin, streptomycin, and fungizone. Five μg of each pJ4Ω expression plasmid was electroporated per 1×10^6 cells.

Given the resemblance of the spliced form of the *erbB*-2 to the sporadic deletions that occur during mammary tumorigenesis in MMTWwild-type *neu* mice, we compared the specific transforming activity of this spliced form to two previously characterized transforming alleles of *erb*B-2. To accomplish this, expression plasmids bearing the wild-type (WT) *erb*B-2 cDNA, a transforming *erb*B-2 allele possessing the point mutation in the transmembrane domain (VE), a previously characterized sporadic deletion discovered in the *neu* proto-oncogene (Siegel *et al.*, 1994) (ECD DEL) or the spliced form of *erb*B-2 (SPLICE) were electroporated into Rat-1 cells and assessed for their capacity to induce transformed foci (Figure 4*a*). Whereas expression of wild-type *erb*B-2 failed to transform Rat-1 fibroblasts, expression of the spliced form and the two previously characterized transforming alleles resulted in the induction of transformed foci (Figure 4*b*). Comparison of the specific transforming activity of the spliced form *erb*B-2 revealed that it transformed at 26% of the level observed with the transmembrane mutant of *erb*B-2 (VE) (Table 1). The spliced form of *erb*B-2 was similar in its transforming activity to the previously characterized sporadic deletion (ECD DEL) (Table 1). Taken together, these observations suggest that the alternative spliced form of *erb*B-2 may represent a causative factor in the induction of human breast tumors.

EXAMPLE 4

Tyrosine Phosphorylation of Altered *erb*B-2 Receptors

Immunoprecipitation and immunblotting

Lysates from established cell lines were prepared from 10 cm tissue culture plates once the cells had reached confluence. The cells were washed twice in ice cold 1XPBS (140 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4) and lysed for 20 min on ice in TNE lysis buffer (50 mM Tris [pH 8.0], 150 mM NaCl, 1% Nonidet P-40, 2 mM EDTA [pH 8.0], 1 mM sodium orthovanadate, 10 mM sodium fluoride, 10 μg/ml Aprotinin, 10 μg/ml Leupeptin). Lysates were cleared by centrifugation for 10 min at 4°C and the protein concentration determined by Bradford assay (Bio-Rad).

Protocols for both immunoprecipitation and immunoblot analysis have been described in detail elsewhere (Muthuswamy *et al.*, 1994). ErbB-2 was immunoprecipitated (Ab-5; 1 μg/600 μg total protein) and blotted (Ab-3; 1:1000) using mouse monoclonal antibodies purchased from Oncogene Research Products, Inc. Phosphotyrosine blots were probed with mouse anti-phosphotyrosine antibodies (PY20; 1:1000) obtained from Transduction Laboratories. Horseradish peroxidase conjugated anti-mouse secondary antibodies (1:2500) (Jackson Laboratories) were used in each case.

To determine whether the oncogenic potential of the spliced form of *erb*B-2 was related to activation of the receptor's catalytic activity, established Rat-1 cells expressing various *erb*B-2 alleles were generated. To assess the state of tyrosine phosphorylation of the receptor, protein extracts derived from the various cell lines were immunoprecipitated with *erb*B-2 specific antibodies and subjected to immunoblot analyses with anti-phosphotyrosine antibodies. As shown in Figure 5a, tyrosine phosphorylated *erb*B-2 could be easily detected in transformed cells expressing the oncogenic alleles of *erb*B-2 but was barely detectable in morphologically normal cell lines expressing elevated levels of wild-type *erb*B-2. A parallel immunoblot for *erb*B-2 revealed that all of the established cell lines expressed similar levels of the receptor (Figure 5b). These observations argue that the oncogenic activation of *erb*B-2 reflects an increase in its catalytic activity, as measured by the extent of tyrosine phosphorylation.

DISCUSSION

The observation that an alternatively spliced form of *erb*B-2 is expressed in a large portion of human breast cancers has important implications in our understanding the molecular basis of this prevalent human malignancy. Consistent with the concept that alternate splicing may play a role in the activation of receptor tyrosine kinases (RTKs), it has been reported that a spliced form of the Ron RTK carries a similar cysteine deletion that leads to its oncogenic activation (Collesi et al, 1996). Moreover, deletions in the EGFR (Sugawa et al., 1990; Wong et al., 1992) and missense mutations in RET (Mulligan et aL, 1993) which also affect cysteine residues situated in the extracellular domains of these RTKs, have frequently been observed in human gliomas and endocrine neoplasias, respectively. In each case, the observed mutations serve to activate the transforming ability of these receptors (Nishikawa et aL, 1994; Santoro et al., 1995; Asai et aL, 1995). Taken together,

these observations suggest that this mechanism of RTK activation may occur frequently in human malignancies. Therapeutic reagents specific to the receptor encoded by this spliced form of *erbB-2 may* have important diagnostic and therapeutic implications in the treatment of human breast cancer.

DETAILED LEGENDS TO FIGURES 3, 4 AND 5

FIGURE 3

Alternate mRNA splicing results in an erbB-2 transcript harboring an in-frame deletion. (a) Schematic representation of a small region of the human *erbB-2* genomic locus indicating the exon-intron boundaries. The values indicated above the schematic represent nucleotide numbers corresponding to the cDNA. Uppercase letters designate coding sequences whereas nuclectides within the introns are in lower case letters. The aberrant splicing event leading to the deletion of an exon (grey) is indicated by the dashed lines. (b) An alignment of the wild-type and alternatively spliced *erbB-2* messages indicating the sequences removed by the aberrant splicing event (grey box). Reverse-transcriptase polymerase chain reaction (RT-PCR) analysis of both normal breast and tumor tissue. The negative control represents an RT-PCR performed in the absence of RNA (distilled water). The PCR was conducted in the presence of [a32P]dCTP and the products resolved on a 5% polyacrylamide gel, dried and exposed to film. The position of the wild-type and spliced products are indicated.

FIGURE 4

Expression of altered ErbB-2 receptors results in morphological transformation. (a) Partial amino acid sequence alignment of the wild-type and altered ErbB-2 receptors tested in Rat-1 focus assays. ErbB-2 (WT) represents the wild-type receptor, ErbB-2 (VE) possesses a valine to glutamic acid point mutation in the transmembrane domain, ErbB-2 (ECD DEL) harbors a deletion which mimics a previously characterized mutation in the Neu receptor (Siegel et aL, 1994; Siegel and Muller, 1996), and ErbB-2 (SPLICE) contains the deletion encoded by the alternatively spliced erbB-2 message. A portion of the transmembrane domain is indicated by the open box and designated TM. The asterisk indicates the position of the valine to glutamic acid substitution in ErbB-2 (VE). (b) Shown are representative plates from focus assay #1 (see Table 1) illustrating the relative transforming abilities of the various forms of ErbB-2 indicated in (a).

FIGURE 5

Altered ErbB-2 receptors are constitutively tyrosine phosphorylated when expressed in Pat-1 cells. (a) Lysates from stable cell lines expressing wild-type ErbB-2 (WT) or the three altered receptors (VE, ECD, SPLICE) were immunoprecipitated for ErbB-2 (Ab5, Oncogene Research Products) and the immunoprecipitates were subjected to immunoblot analysis with phosphotyrosine-specific antibodies (PY20, Transduction Laboratories). (b) Lysates from stable cell lines expressing the various ErbB-2 receptors were immunoprecipitated for ErbB-2 (Ab5, Oncogene Research Products) and the immunoprecipitates were subjected to immunoblot analysis with ErbB-2/Neu-specific antibodies (AW, Oncogene Research Products). In each panel, Ratl indicates the parental cell line which does not contain any of the erbB-2 expression plasmids. The position of ErbB-2 is indicated in each panel by the arrow.

Having illustrated and described the principles of the invention in a preferred embodiment, it should be appreciated to those skilled in the art that the invention can be modified in arrangement and detail without departure from such principles. We claim all modifications coming within the scope of the following claims.

All publications, patents and patent applications referred to herein are incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

FULL CITATIONS FOR REFERENCES REFERRED TO IN THE SPECIFICATION

1. **Asai, N., T. Washita, M. Matsuyama and M. Takahashi.** 1995. Mechanism of activation of the ret proto-oncogene by multiple endocrine neoplasia 2A mutations. Mol. Cell. Biol. 15:1613-1619.

2. **Collesi, C., M. M. Santoro, G. Gaudino and P. M. Comoglio.** 1996. A splicing variant of the *RON* transcript induces constitutive tyrosine kinase activity and an invasive phenotype. Mol. Cell. Biol. 16:5518-5526.

3. **Guy, C. T., R. D. Cardiff and W. J. Muller.** 1996. Activated *neu* induces rapid tumor progression. I Biol. Chem. 271:7673-7678.

4. **Guy, C. T., M. A. Webster, M. Schaller, T. J. Parsons, R. D. Cardiff and W. J. Muller.** 1992. Expression of the *neu* protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. Proc. Natl. Aced. Sci. USA 89:1057810582.

5. **Hynes, N. E. and D. F. Stern.** 1994. The biology of *erbB-2/neulHER-2* and its role in cancer. Biochim. Biophys. Acta 1198:165-184.

6. **Muller, W. J., E. Sinn, P. K. Pattengale, R. Wallace and P. Leder.** 1988. Single step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell 54:105-115.

7. **Mulligan, L. M., J. B. J. Kwok, C. S. Healey, M. I Elsdon, C. Eng, E. Gardner, D. R. Love, S. E. Mole, J. K. Moore, L. Papi, M. A. Ponder, H. Telenlus, A. Tunnacliffe and B. A. J. Ponder.** 1993. Germ-line mutations of the PETproto-oncogene in multiple endocrine neoplasia type 2A. Nature 363:458-460.

8. **Nishikawa, R., X.-D. Jl, R. C. Harmon, C. S. Lazar, G. N. Gill, W. K. Cavanee and H.-J. Su Huang.** 1994. A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity. Proc. Nall. Acad. Sci. USA 91:7727-7731.

9. **Santoro, M., F. Carlomagno, A. Romano, D. P. Bottaro, N. A. Dathan, M. Grieco, A. Fusco, G. Vecchio, B. Matoskova, M. H. Kraus and P. P. DiFlore.** 1995. Activation of *RETas a* dominant transforming gene by germline mutations of MEN2A and MEN213. Science 267:381-383.

10. **Siegel, P. M., D. L. Dankort, W. R. Hardy and W. J. Muller.** 1994. Novel activating mutations in the *neu* proto-oncogene involved in induction of mammary tumors. Mol. Cell. Biol. 14:7068-7077.

11. **Siegel, P. M. and W. J. Muller.** 1996. Mutations affecting conserved cysteine residues within the extracellular domain of Neu promote receptor dimerization and activation. Proc. Natl. Acad. Sci. USA 93:8878-8883.

12. **Slamon, D. J., G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich and W. L. McGuire.** 1987. Human breast cancer: correlation of relapse and survival with the amplification of the HER-21neu oncogene. Science 235:177-182.

13. Slamon, D. J., W. Godolphin, L. A. Jones, J. A. Holt, S. G. Wong, D. E. Keith, W. J. Levin, S. G. Stuart, J. Udove, A. Ulirich and M. F. Press. 1989. Studies of the HER-21neu proto-oncogene in human breast and ovarian cancer. Science 244:707-712.

14. Sugawa, N., A. J. Ekstrand, C. D. James and V. P. Collins. 1990. Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human gliomas. Proc. Natl. Acad. Sci. USA 87:8602-8606.

15. Wong, A. J., J. M. Ruppert, S. H. Bigner, C. H. Grzeschik, P. A. Humphrey, D. S. Bigner and B. Vogelstein. 1992. Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc. Nati. Acad. Sci. USA 89:2965-2969.

16. Coussens, L., Yang-Feng, T.L., Liao, Y.-C., Chen, E., Gray, A., McGrath, J., Seeburg, P.H., Libermann, T.A., Schlessinger, J., Francke, U., Levinson, A. and Ullrich, A. (1985) Tyrosine Kinase Receptor with Extensive Homology to EGF Receptor Shares Chromosomal Location with *neu* Oncogene. *Science*, 230, 1132-1139.

17. Muthuswamy, S.K., Siegel, P.M., Dankort, D.L., Webster, M.A. and Muller, W.J. (1994) Mammary tumors expressing the *neu* proto-oncogene possess elevated c-Src tyrosine kinase activity. *Mol. Cell. Biol.*, 14, 735-743.

Table 1. Transformation of Rat-1 cells by altered ErbB-2 receptors.

Expression Plasmid[a]	Focus Assay #1		Focus Assay #2		Focus Assay #3		Relative Transforming ability[d] (%)
	Average No. of foci/plate[b]	% Transform. of pJ4ΩB2(VE)[c]	Average No. of foci/plate[b]	% Transform. of pJ4ΩB2(VE)[c]	Average No. of foci/plate[b]	% Transform. of pJ4ΩB2(VE)[c]	
pJ4ΩB2 (WT)	0	0	0	0	0	0	0
pJ4ΩB2 (VE)	284±15	100	523±28	100	672±13	100	100±0
pJ4ΩB2 (ECD DEL)	63±6	22	130±16	25	195±14	29	25±4
pJ4ΩB2 (SPLICE)	73±8	26	127±15	24	194±14	29	26±3

Three independent focus assays were performed with established Rat-1 fibroblasts. The first two experiments were conducted with plasmid DNA from the same large-scale preparation, while the third experiment was performed with DNA obtained from an independent plasmid preparation. The Rat-1 fibroblasts were maintained in a monolayer for 14 days following electroporation at which time they were stained with Geimsa.

[a] All cDNAs placed under the transcriptional control of the Moloney murine leukemia virus long terminal repeat (Mo-MuLV LTR).

[b] Values listed represent the mean number of foci counted on 8 plates ± the standard deviations.

[c] Values represent the ratio of the mean number of foci obtained for each constuct with respect to pJ4ΩB2(VE).

[d] Values listed represent the mean transforming abilities from all three experiments ± the standard deviations.

WE CLAIM:

1. A purified and isolated nucleic acid molecule comprising a sequence encoding an *erb*B-2 receptor protein with an in frame deletion of 16 amino acids two of which are cysteine residues (SPLICE *erb*B-2 protein).

2. A purified and isolated nucleic acid molecule according to claim 1 wherein the protein is constitutively tyrosine phosphorylated.

3. A purified and isolated nucleic acid molecule as claimed in claim 1, comprising (i) a nucleic acid sequence encoding a SPLICE *erb*B-2 protein having the amino acid sequence according to Seq.ID.No.2 and, (ii) nucleic acid sequences complementary to (i).

4. A purified and isolated nucleic acid molecule as claimed in claim 1, comprising

(i) a nucleic acid sequence encoding a SPLICE *erb*B-2 protein having the nucleic acid sequence according to Seq. ID.No.1 wherein T can also be U;

(ii) a nucleic acid sequence complementary to (i);

(iii) a nucleic acid molecule differing from any of the nucleic acids of (i) and (ii) in codon sequences due to the degeneracy of the genetic code.

5. A purified and isolated nucleic acid molecule comprising a sequence which hybridizes to the nucleic acid molecule as claimed in claim 3.

6. A recombinant expression vector adapted for transformation of a host cell comprising a nucleic acid molecule as claimed in claim 1 and one or more transcription and translation elements operatively linked to the nucleic acid molecule.

7. A host cell containing a recombinant expression vector as claimed in claim 6.

8. A method for preparing a SPLICE *erb*B-2 protein comprising (a) transferring a recombinant expression vector as claimed in claim 6 into a host cell; (b) selecting transformed host cells from untransformed host cells; (c) culturing a selected transformed host cell under conditions which allow expression of the protein; and (d) isolating the protein.

9. A purified and isolated SPLICE *erb*B-2 protein.

10. The purified and isolated protein according to claim 9 wherein the protein is constitutively tyrosine phosphorylated.

11. A purified and isolated protein as claimed in claim 9 or 10, which has an amino acid sequence according to Seq.ID.No. 2, or a fragment, analog or derivative thereof.

12. A method of modulating cell transformation comprising administering an effective amount of a SPLICE *erb*B-2 protein or a nucleic acid encoding a SPLICE *erb*B-2 protein to a cell or animal.

13. A method according to claim 12 wherein the protein has the amino acid sequence according to Seq.ID.NO 2 or a fragment thereof.

14. A method according to claim 12 or 13 wherein the nucleic acid molecule encoding the protein has the sequence according to Seq.ID.NO.1.

15. A method of inhibiting or reducing cell transformation comprising administering an antisense molecule that is complimentary to a nucleic acid molecule encoding a SPLICE *erb*B-2 protein.

16. The method of claim 15 wherein the cell transformation occurs in breast tissue.

17. A method of modulating SPLICE *erb*B-2 activity comprising administering an effective amount of inhibitor of SPLICE *erb*B-2 protein activity to a cell or animal in need thereof.

18. A method according to claim 17 wherein the inhibitor is selected from the group consisting of antibodies to the SPLICE *erb*B-2 protein and antisense oligonucleotides complimentary to a portion of a nucleic acid sequence encoding the SPLICE *erb*B-2 protein.

19. A method according to claim 18 wherein the modulation is inhibition of cancer cell growth in breast tumors.

20. A method for identifying a substance which is capable of binding to a SPLICE *erb*B-2 protein or inactive form thereof, comprising the following steps:

(i) reacting the SPLICE *erb*B-2 protein, or inactivated form, with at least one substance which potentially can bind with the SPLICE *erb*B-2 protein or an inactivated form thereof, under conditions which permit the formation of complexes between the substanc and the protein or an inactivated form thereof; and

(ii) assaying for complexes, for free substance, for non-complexed SPLICE *erb*B-2 protein or for activation of SPLICE *erb*B-2

21. A method for assaying a medium for the presence of an agonist or antagonist of the interaction of a SPLICE *erb*B-2 protein or an activated form thereof, and a substance which binds to the protein or an activated form thereof comprising the steps of:

(i) combining a known concentration of a SPLICE *erb*B-2 protein, with a substance which is capable of binding to the protein and a suspected agonist or antagonist substance under conditions which permit the formation of complexes between the substance and SPLICE *erb*B-2 protein;

(ii) assaying for complexes, free substance, non-complexed SPLICE *erb*B-2 protein, or for activation of the protein.

FIGURE 1

aaggggaggtaaccctggcccctttggtcggggccccgggcagccgcgcgccccttccca
cggggccctttactgcgccgcgcgcccggcccccacccctcgcagcaccccgcgccccgc
gccctcccagccgggtccagccggagccatggggccggagccgcagtgagcacc**ATG**GAG
CTGGCGGCCTTGTGCCGCTGGGGGCTCCTCCTCGCCCTCTTGCCCCCCGGAGCCGCGAGC
ACCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCAC
CTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACTC
ACCTACCTGCCCACCAATGCCAGCCTGTCCTTCCTGCAGGATATCCAGGAGGTGCAGGGC
TACGTGCTCATCGCTCACAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTGTG
CGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGAGACCCG
CTGAACAATACCACCCCTGTCACAGGGGCCTCCCCAGGAGGCCTGCGGGAGCTGCAGCTT
CGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAGCTCTGC
TACCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACCAGCTGGCTCTCACA
CTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAGGGCTCC
CGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGTGCCGGT
GGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGTGCTGCC
GGCTGCACGGGCCCCAAGCACTCTGACTGCCTGGCCTGCCTCCACTTCAACCACAGTGGC
ATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAACACAGACACGTTTGAGTCCATG
CCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCCTACAAC
TACCTTTCTACGGACGTGGGATCCTGCACCCTCGTCTGCCCCCTGCACAACCAAGAGGTG
ACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGAGTGTGC
TATGGTCTGGGCATGGAGCACTTGCGAGAGGTGAGGGCAGTTACCAGTGCCAATATCCAG
GAGTTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGCTTTGAT
GGGGACCCAGCCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTTGAGACT
CTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACCTC
AGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCTGCACAATGGCGCCTACTCG
CTGACCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACTGGGC
AGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTGCCCTGG
GACCAGCTCTTTCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGGAC
GAGTGTGTGGGCGAGGGCCTGGCCTGCCACCAGCTGTGCGCCCGAGGGCACTGCTGGGGT
CCAGGGCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGCGTGGAG
GAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGTTTGCCG

FIGURE 1 (cont'd)

TGCCACCCTGAGTGTCAGCCCCAGAATGGCTCAGTGACCTGTTTGGACCGGAGGCTGAC
CAGTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCCAGC
GGTGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCA
TGCCAGCCTTGCCCCATCAACTGCACCCAC**TC**----------------------------
--------------------CCCTCTGACGTCCATCATCTCTGCGGTGGTTGGCATTCTG
CTGGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAGAAGATC
CGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTGACACCT
AGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTGAGGAAG
GTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGAT
GGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTTGAGGGAAAACACATCCCCCAAA
GCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTCCCCATATGTC
TCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTTATGCCC
TATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCTGGGCTCCCAGGACCTG
CTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGGCTCGTA
CACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAAATTACA
GACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGCAGATGGGGGC
AAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTTCACCCACCAG
AGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCCAAACCT
TACGATGGGATCCCAGCCCGGGAGATCCCTGACCTGCTGGAAAAGGGGGAGCGGCTGCCC
CAGCCCCCCATCTGCACCATTGATGTCTACATGATCATGGTCAAATGTTGGATGATTGAC
TCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCATGGCCAGGGAC
CCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCAGCCAGTCCCTTGGACAGC
ACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAG
TATCTGGTACCCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCCGGGCGCTGGGGGCATG
GTCCACCACAGGCACCGCAGCTCATCTACCAGGAGTGGCGGTGGGGACCTGACACTAGGG
CTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCCTCCGAAGGGGCTGGC
TCCGATGTATTTGATGGTGACCTGGGAATGGGGGCAGCCAAGGGGCTGCAAAGCCTCCCC
ACACATGACCCCAGCCCTCTACAGCGGTACAGTGAGGACCCCACAGTACCCCTGCCCTCT
GAGACTGATGGCTACGTTGCCCCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCAG
CCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCCTGCTGCCCGACCT

FIGURE 1 (cont'd)

GCTGGTGCCACTCTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGGGGTCGTCAAA
GACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCCAGGGAGGA
GCTGCCCCTCAGCCCCACCCTCCTCCTGCCTTCAGCCCAGCCTTCGACAACCTCTATTAC
TGGGACCAGGACCCACCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGACACCTACG
GCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG**TGA**accagaaggccaagtccg
cagaagccctgatgtgtcctcagggagcagggaaggcctgacttctgctggcatcaagag
gtgggagggccctccgaccacttccaggggaacctgccatgccaggaacctgtcctaagg
aaccttccttcctgcttgagttcccagatggctggaaggggtccagcctcgttggaagag
gaacagcactggggagtctttgtggattctgaggccctgcccaatgagactctagggtcc
agtggatgccacagcccagcttggccctttccttccagatcctgggtactgaaagcctta
gggaagctggcctgagaggggaagcggccctaagggagtgtctaagaacaaaagcgaccc
attcagagactgtccctgaaacctagtactgccccccatgaggaaggaacagcaatggtg
tcagtatccaggctttgtacagagtgcttttctgtttagtttttactttttttgttttgt
tttttaaagatgaaataaagacccaggggggag

FIGURE 2

M E

L A A L C R W G L L L A L L P P G A A S

T Q V C T G T D M K L R L P A S P E T H

L D M L R H L Y Q G C Q V V Q G N L E L

T Y L P T N A S L S F L Q D I Q E V Q G

Y V L I A H N Q V R Q V P L Q R L R I V

R G T Q L F E D N Y A L A V L D N G D P

L N N T T P V T G A S P G G L R E L Q L

R S L T E I L K G G V L I Q R N P Q L C

Y Q D T I L W K D I F H K N N Q L A L T

L I D T N R S R A C H P C S P M C K G S

R C W G E S S E D C Q S L T R T V C A G

G C A R C K G P L P T D C C H E Q C A A

G C T G P K H S D C L A C L H F N H S G

I C E L H C P A L V T Y N T D T F E S M

P N P E G R Y T F G A S C V T A C P Y N

Y L S T D V G S C T L V C P L H N Q E V

T A E D G T Q R C E K C S K P C A R V C

Y G L G M E H L R E V R A V T S A N I Q

E F A G C K K I F G S L A F L P E S F D

G D P A S N T A P L Q P E Q L Q V F E T

L E E I T G Y L Y I S A W P D S L P D L

S V F Q N L Q V I R G R I L H N G A Y S

L T L Q G L G I S W L G L R S L R E L G

S G L A L I H H N T H L C F V H T V P W

D Q L F R N P H Q A L L H T A N R P E D

E C V G E G L A C H Q L C A R G H C W G

P G P T Q C V N C S Q F L R G Q E C V E

E C R V L Q G L P R E Y V N A R H C L P

C H P E C Q P Q N G S V T C F G P E A D

Q C V A C A H Y K D P P F C V A R C P S

G V K P D L S Y M P I W K F P D E E G A

FIGURE 2 (cont'd)

C Q P C P I N C T H S - - - - - - - - -

- - - - - - - P L T S I I S A V V G I L

L V V V L G V V F G I L I K R R Q Q K I

R K Y T M R R L L Q E T E L V E P L T P

S G A M P N Q A Q M R I L K E T E L R K

V K V L G S G A F G T V Y K G I W I P D

G E N V K I P V A I K V L R E N T S P K

A N K E I L D E A Y V M A G V G S P Y V

S R L L G I C L T S T V Q L V T Q L M P

Y G C L L D H V R E N R G R L G S Q D L

L N W C M Q I A K G M S Y L E D V R L V

H R D L A A R N V L V K S P N H V K I T

D F G L A R L L D I D E T E Y H A D G G

K V P I K W M A L E S I L R R R F T H Q

S D V W S Y G V T V W E L M T F G A K P

Y D G I P A R E I P D L L E K G E R L P

Q P P I C T I D V Y M I M V K C W M I D

S E C R P R F R E L V S E F S R M A R D

P Q R F V V I Q N E D L G P A S P L D S

T F Y R S L L E D D D M G D L V D A E E

Y L V P Q Q G F F C P D P A P G A G G M

V H H R H R S S S T R S G G G D L T L G

L E P S E E E A P R S P L A P S E G A G

S D V F D G D L G M G A A K G L Q S L P

T H D P S P L Q R Y S E D P T V P L P S

E T D G Y V A P L T C S P Q P E Y V N Q

P D V R P Q P P S P R E G P L P A A R P

A G A T L E R P K T L S P G K N G V V K

D V F A F G G A V E N P E Y L T P Q G G

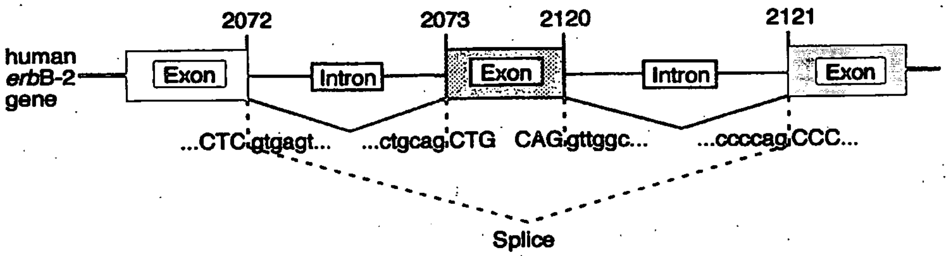
A A P Q P H P P P A F S P A F D N L Y Y

W D Q D P P E R G A P P S T F K G T P T

A E N P E Y L G L D V P V *

FIGURE 3

A



B

WT (Human *erb*B-2)

ACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTG ACGTCC
T H S C V D L D D K G C P A E Q R A S P L T S

Alternate splice (Human *erb*B-2)

ACCCACTC CCCTCTGACGTCC
T H S P L T S

C

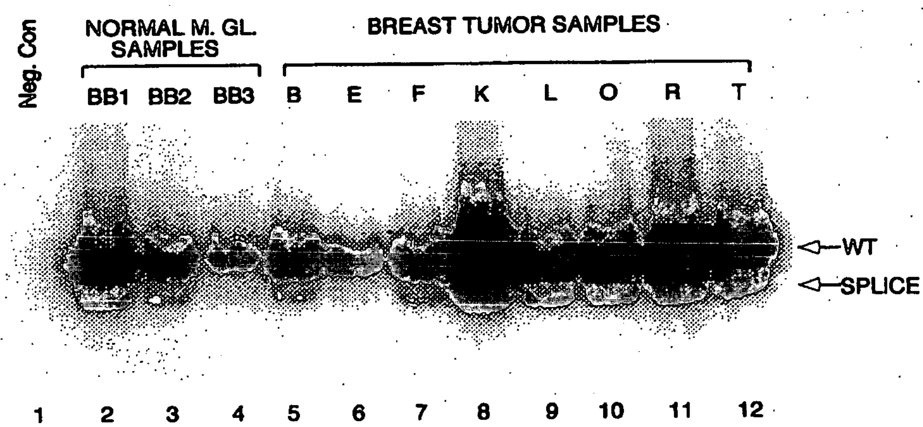


FIGURE 4

a

```
                                                                      TM
ErbB-2(WT)        :CQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVVGILLV
                                                                       *
ErbB-2(VE)        :CQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVEGILLV

ErbB-2(ECD DEL)   :CQPCPINCTHSCV                    ASPLTSIISAVVGILLV

ErbB-2(SPLICE)    :CQPCPINCTHS                         PLTSIISAVVGILLV
```

b

WT VE ECD DEL SPLICE

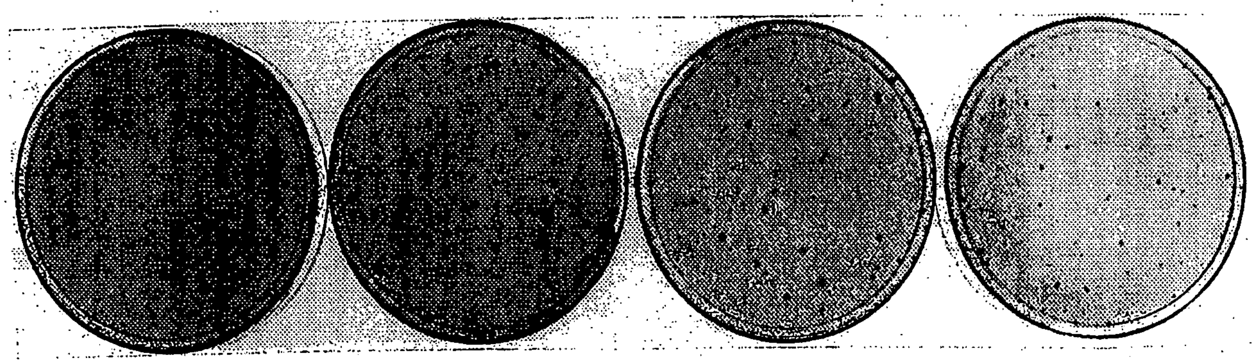
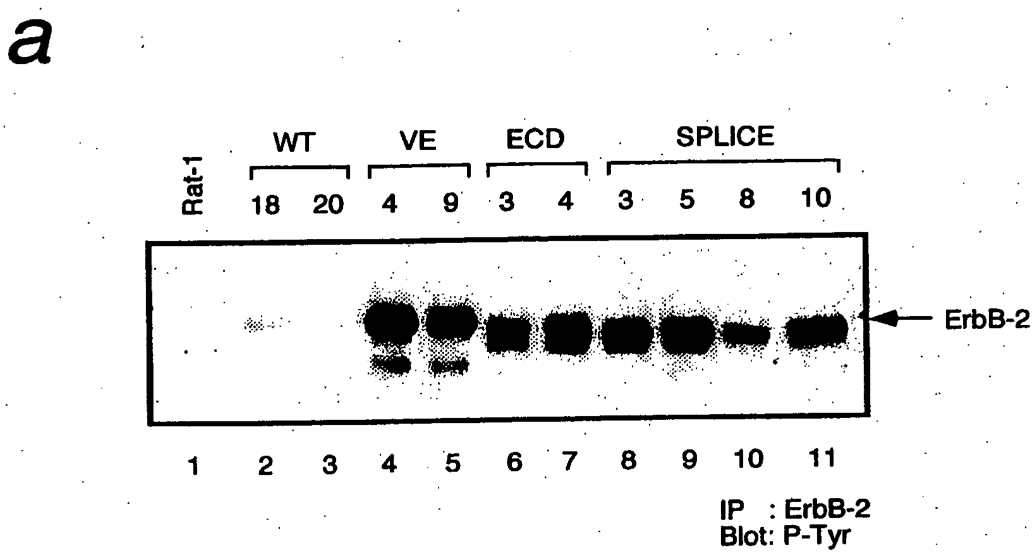
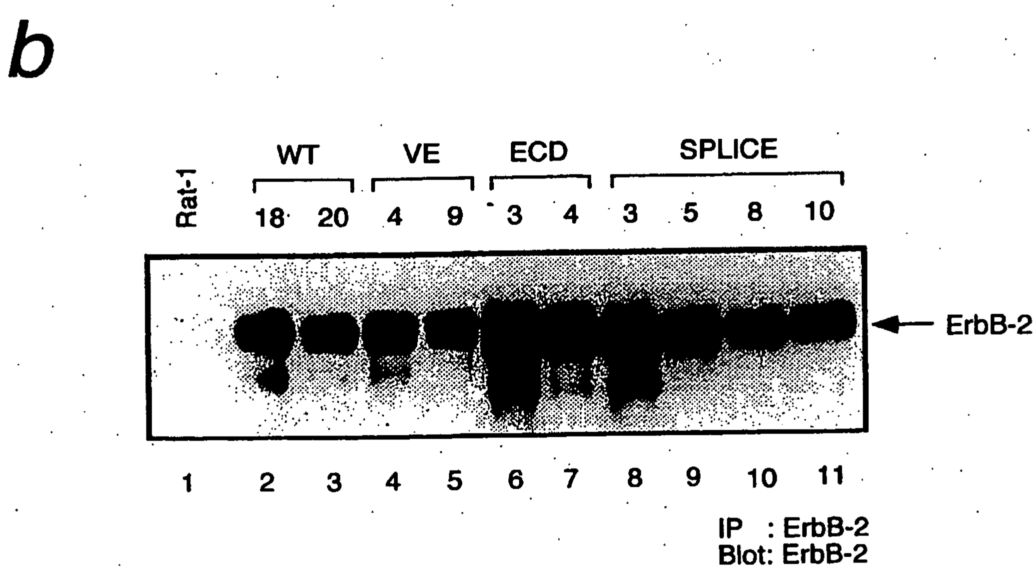


FIGURE 5


a
WT
VE
ECD
SPLICE
Rat-1
18
20
4
9
3
4
3
5
8
10
ErbB-2
1
2
3
4
5
6
7
8
9
10
11
IP : ErbB-2
Blot: P-Tyr


b
WT
VE
ECD
SPLICE
Rat-1
18
20
4
9
3
4
3
5
8
10
ErbB-2
1
2
3
4
5
6
7
8
9
10
11
IP : ErbB-2
Blot: ErbB-2

SEQUENCE LISTING

<110> Muller, William J.
Siegel, Peter M.

<120> Novel Spliced Form of ErbB-2/Neu Oncogene

<130> 3244-012

<140>
<141>

<160> 12

<170> PatentIn Ver. 2.0

<210> 1
<211> 4472
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (175)..(3942)

<400> 1
aaggggaggt aaccctggcc cctttggtcg gggccccggg cagccgcgcg cccttccca 60

cggggccctt tactgcgccg cgcgcccggc ccccacccct cgcagcaccc cgcgccccgc 120

gccctcccag ccgggtccag ccggagccat ggggccggag ccgcagtgag cacc atg 177
Met
1

gag ctg gcg gcc ttg tgc cgc tgg ggg ctc ctc ctc gcc ctc ttg ccc 225
Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu Pro
5 10 15

ccc gga gcc gcg agc acc caa gtg tgc acc ggc aca gac atg aag ctg 273
Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys Leu
20 25 30

cgg ctc cct gcc agt ccc gag acc cac ctg gac atg ctc cgc cac ctc 321
Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His Leu
35 40 45

tac cag ggc tgc cag gtg gtg cag gga aac ctg gaa ctc acc tac ctg 369
Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr Leu
50 55 60 65

ccc acc aat gcc agc ctg tcc ttc ctg cag gat atc cag gag gtg cag 417
Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val Gln
70 75 80

ggc tac gtg ctc atc gct cac aac caa gtg agg cag gtc cca ctg cag 465
Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu Gln
85 90 95

agg ctg cgg att gtg cga ggc acc cag ctc ttt gag gac aac tat gcc 513
Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr Ala
100 105 110

ctg gcc gtg cta gac aat gga gac ccg ctg aac aat acc acc cct gtc 561
Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro Val
115 120 125

```
aca ggg gcc tcc cca gga ggc ctg cgg gag ctg cag ctt cga agc ctc    609
Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser Leu
130                 135                 140                 145

aca gag atc ttg aaa gga ggg gtc ttg atc cag cgg aac ccc cag ctc    657
Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln Leu
                150                 155                 160

tgc tac cag gac acg att ttg tgg aag gac atc ttc cac aag aac aac    705
Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn Asn
            165                 170                 175

cag ctg gct ctc aca ctg ata gac acc aac cgc tct cgg gcc tgc cac    753
Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys His
        180                 185                 190

ccc tgt tct ccg atg tgt aag ggc tcc cgc tgc tgg gga gag agt tct    801
Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser Ser
    195                 200                 205

gag gat tgt cag agc ctg acg cgc act gtc tgt gcc ggt ggc tgt gcc    849
Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys Ala
210                 215                 220                 225

cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat gag cag tgt gct    897
Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys Ala
                230                 235                 240

gcc ggc tgc acg ggc ccc aag cac tct gac tgc ctg gcc tgc ctc cac    945
Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu His
            245                 250                 255

ttc aac cac agt ggc atc tgt gag ctg cac tgc cca gcc ctg gtc acc    993
Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val Thr
        260                 265                 270

tac aac aca gac acg ttt gag tcc atg ccc aat ccc gag ggc cgg tat    1041
Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg Tyr
    275                 280                 285

aca ttc ggc gcc agc tgt gtg act gcc tgt ccc tac aac tac ctt tct    1089
Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu Ser
290                 295                 300                 305

acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg cac aac caa gag    1137
Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln Glu
                310                 315                 320

gtg aca gca gag gat gga aca cag cgg tgt gag aag tgc agc aag ccc    1185
Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys Pro
            325                 330                 335

tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac ttg cga gag gtg    1233
Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu Val
        340                 345                 350

agg gca gtt acc agt gcc aat atc cag gag ttt gct ggc tgc aag aag    1281
Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys Lys
    355                 360                 365

atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt gat ggg gac cca    1329
Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp Pro
370                 375                 380                 385

gcc tcc aac act gcc ccg ctc cag cca gag cag ctc caa gtg ttt gag    1377
Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe Glu
                390                 395                 400
```

```
act ctg gaa gag atc aca ggt tac cta tac atc tca gca tgg ccg gac    1425
Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro Asp
            405                 410                 415

agc ctg cct gac ctc agc gtc ttc cag aac ctg caa gta atc cgg gga    1473
Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg Gly
        420                 425                 430

cga att ctg cac aat ggc gcc tac tcg ctg acc ctg caa ggg ctg ggc    1521
Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu Gly
    435                 440                 445

atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg ggc agt gga ctg    1569
Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly Leu
450                 455                 460                 465

gcc ctc atc cac cat aac acc cac ctc tgc ttc gtg cac acg gtg ccc    1617
Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val Pro
                470                 475                 480

tgg gac cag ctc ttt cgg aac ccg cac caa gct ctg ctc cac act gcc    1665
Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr Ala
            485                 490                 495

aac cgg cca gag gac gag tgt gtg ggc gag ggc ctg gcc tgc cac cag    1713
Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His Gln
        500                 505                 510

ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc acc cag tgt gtc    1761
Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys Val
    515                 520                 525

aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg gag gaa tgc cga    1809
Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys Arg
530                 535                 540                 545

gta ctg cag ggg ctc ccc agg gag tat gtg aat gcc agg cac tgt ttg    1857
Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys Leu
                550                 555                 560

ccg tgc cac cct gag tgt cag ccc cag aat ggc tca gtg acc tgt ttt    1905
Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys Phe
            565                 570                 575

gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac tat aag gac cct    1953
Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp Pro
        580                 585                 590

ccc ttc tgc gtg gcc cgc tgc ccc agc ggt gtg aaa cct gac ctc tcc    2001
Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu Ser
    595                 600                 605

tac atg ccc atc tgg aag ttt cca gat gag gag ggc gca tgc cag cct    2049
Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln Pro
610                 615                 620                 625

tgc ccc atc aac tgc acc cac tcc tgt gtg gac ctg gat gac aag ggc    2097
Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly
                630                 635                 640

tgc ccc gcc gag cag aga gcc agc cct ctg acg tcc atc atc tct gcg    2145
Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser Ala
            645                 650                 655

gtg gtt ggc att ctg ctg gtc gtg gtc ttg ggg gtg gtc ttt ggg atc    2193
Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly Ile
        660                 665                 670
```

```
ctc atc aag cga cgg cag cag aag atc cgg aag tac acg atg cgg aga    2241
Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg Arg
    675                 680                 685

ctg ctg cag gaa acg gag ctg gtg gag ccg ctg aca cct agc gga gcg    2289
Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Ala
690                 695                 700                 705

atg ccc aac cag gcg cag atg cgg atc ctg aaa gag acg gag ctg agg    2337
Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu Arg
                710                 715                 720

aag gtg aag gtg ctt gga tct ggc gct ttt ggc aca gtc tac aag ggc    2385
Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly
            725                 730                 735

atc tgg atc cct gat ggg gag aat gtg aaa att cca gtg gcc atc aaa    2433
Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile Lys
        740                 745                 750

gtg ttg agg gaa aac aca tcc ccc aaa gcc aac aaa gaa atc tta gac    2481
Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp
    755                 760                 765

gaa gca tac gtg atg gct ggt gtg ggc tcc cca tat gtc tcc cgc ctt    2529
Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg Leu
770                 775                 780                 785

ctg ggc atc tgc ctg aca tcc acg gtg cag ctg gtg aca cag ctt atg    2577
Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu Met
                790                 795                 800

ccc tat ggc tgc ctc tta gac cat gtc cgg gaa aac cgc gga cgc ctg    2625
Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg Leu
            805                 810                 815

ggc tcc cag gac ctg ctg aac tgg tgt atg cag att gcc aag ggg atg    2673
Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly Met
        820                 825                 830

agc tac ctg gag gat gtg cgg ctc gta cac agg gac ttg gcc gct cgg    2721
Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala Arg
    835                 840                 845

aac gtg ctg gtc aag agt ccc aac cat gtc aaa att aca gac ttc ggg    2769
Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly
850                 855                 860                 865

ctg gct cgg ctg ctg gac att gac gag aca gag tac cat gca gat ggg    2817
Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp Gly
                870                 875                 880

ggc aag gtg ccc atc aag tgg atg gcg ctg gag tcc att ctc cgc cgg    2865
Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg Arg
            885                 890                 895

cgg ttc acc cac cag agt gat gtg tgg agt tat ggt gtg act gtg tgg    2913
Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp
        900                 905                 910

gag ctg atg act ttt ggg gcc aaa cct tac gat ggg atc cca gcc cgg    2961
Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala Arg
    915                 920                 925

gag atc cct gac ctg ctg gaa aag ggg gag cgg ctg ccc cag ccc ccc    3009
Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro
930                 935                 940                 945
```

```
atc tgc acc att gat gtc tac atg atc atg gtc aaa tgt tgg atg att    3057
Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile
                950                 955                 960

gac tct gaa tgt cgg cca aga ttc cgg gag ttg gtg tct gaa ttc tcc    3105
Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe Ser
            965                 970                 975

cgc atg gcc agg gac ccc cag cgc ttt gtg gtc atc cag aat gag gac    3153
Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu Asp
        980                 985                 990

ttg ggc cca gcc agt ccc ttg gac agc acc ttc tac cgc tca ctg ctg    3201
Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu Leu
    995                1000                1005

gag gac gat gac atg ggg gac ctg gtg gat gct gag gag tat ctg gta    3249
Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu Val
1010                1015                1020                1025

ccc cag cag ggc ttc ttc tgt cca gac cct gcc ccg ggc gct ggg ggc    3297
Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly Gly
               1030                1035                1040

atg gtc cac cac agg cac cgc agc tca tct acc agg agt ggc ggt ggg    3345
Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly Gly
           1045                1050                1055

gac ctg aca cta ggg ctg gag ccc tct gaa gag gag gcc ccc agg tct    3393
Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg Ser
       1060                1065                1070

cca ctg gca ccc tcc gaa ggg gct ggc tcc gat gta ttt gat ggt gac    3441
Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly Asp
   1075                1080                1085

ctg gga atg ggg gca gcc aag ggg ctg caa agc ctc ccc aca cat gac    3489
Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His Asp
1090                1095                1100                1105

ccc agc cct cta cag cgg tac agt gag gac ccc aca gta ccc ctg ccc    3537
Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu Pro
               1110                1115                1120

tct gag act gat ggc tac gtt gcc ccc ctg acc tgc agc ccc cag cct    3585
Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln Pro
           1125                1130                1135

gaa tat gtg aac cag cca gat gtt cgg ccc cag ccc cct tcg ccc cga    3633
Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro Arg
       1140                1145                1150

gag ggc cct ctg cct gct gcc cga cct gct ggt gcc act ctg gaa agg    3681
Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg
   1155                1160                1165

ccc aag act ctc tcc cca ggg aag aat ggg gtc gtc aaa gac gtt ttt    3729
Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe
1170                1175                1180                1185

gcc ttt ggg ggt gcc gtg gag aac ccc gag tac ttg aca ccc cag gga    3777
Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly
               1190                1195                1200

gga gct gcc cct cag ccc cac cct cct cct gcc ttc agc cca gcc ttc    3825
Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe
           1205                1210                1215
```

```
gac aac ctc tat tac tgg gac cag gac cca cca gag cgg ggg gct cca    3873
Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro
     1220                1225                1230

ccc agc acc ttc aaa ggg aca cct acg gca gag aac cca gag tac ctg    3921
Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr Leu
 1235                1240                1245

ggt ctg gac gtg cca gtg tga accagaaggc caagtcccag aagccctgat      3972
Gly Leu Asp Val Pro Val
1250                1255

gtgtcctcag ggagcaggga aggcctgact tctgctggca tcaagaggtg ggagggccct 4032

ccgaccactt ccaggggaac ctgccatgcc aggaacctgt cctaaggaac cttccttcct 4092

gcttgagttc ccagatggct ggaaggggtc cagcctcgtt ggaagaggaa cagcactggg 4152

gagtctttgt ggattctgag gccctgccca atgagactct agggtccagt ggatgccaca 4212

gcccagcttg gccctttcct tccagatcct gggtactgaa agccttaggg aagctggcct 4272

gagaggggaa gcggccctaa gggagtgtct aagaacaaaa gcgacccatt cagagactgt 4332

ccctgaaacc tagtactgcc ccccatgagg aaggaacagc aatggtgtca gtatccaggc 4392

tttgtacaga gtgcttttct gtttagtttt tactttttttt gttttgtttt tttaaagatg 4452

aaataaagac ccagggggag                                            4472


<210> 2
<211> 1255
<212> PRT
<213> Homo sapiens

<400> 2
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                   10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160
```

```
Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
                325                 330                 335

Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
            340                 345                 350

Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
        355                 360                 365

Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
    370                 375                 380

Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385                 390                 395                 400

Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
                405                 410                 415

Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
            420                 425                 430

Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
        435                 440                 445

Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
    450                 455                 460

Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465                 470                 475                 480

Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
                485                 490                 495

Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
            500                 505                 510

Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
```

```
        515                 520                 525

Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
    530                 535                 540

Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545                 550                 555                 560

Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
                565                 570                 575

Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
            580                 585                 590

Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
        595                 600                 605

Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
    610                 615                 620

Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
625                 630                 635                 640

Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser
                645                 650                 655

Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
            660                 665                 670

Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
        675                 680                 685

Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
    690                 695                 700

Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
705                 710                 715                 720

Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
                725                 730                 735

Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
            740                 745                 750

Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
        755                 760                 765

Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
    770                 775                 780

Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
785                 790                 795                 800

Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
                805                 810                 815

Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
            820                 825                 830

Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
        835                 840                 845

Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
    850                 855                 860

Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865                 870                 875                 880
```

Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
885 890 895

Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
900 905 910

Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
915 920 925

Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
930 935 940

Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945 950 955 960

Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
965 970 975

Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
980 985 990

Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
995 1000 1005

Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu
1010 1015 1020

Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala Gly
025 1030 1035 1040

Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly
1045 1050 1055

Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala Pro Arg
1060 1065 1070

Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp Val Phe Asp Gly
1075 1080 1085

Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln Ser Leu Pro Thr His
1090 1095 1100

Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro Leu
105 1110 1115 1120

Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu Thr Cys Ser Pro Gln
1125 1130 1135

Pro Glu Tyr Val Asn Gln Pro Asp Val Arg Pro Gln Pro Pro Ser Pro
1140 1145 1150

Arg Glu Gly Pro Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu
1155 1160 1165

Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val
1170 1175 1180

Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln
185 1190 1195 1200

Gly Gly Ala Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala
1205 1210 1215

Phe Asp Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala
1220 1225 1230

Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr

1235 1240 1245

Leu Gly Leu Asp Val Pro Val
1250 1255

<210> 3
<211> 18
<212> DNA
<213> Homo sapiens

<400> 3
tttccagatg aggagggc 18

<210> 4
<211> 27
<212> DNA
<213> Homo sapiens

<400> 4
cggaattcct gtcaccagct gcaccgt 27

<210> 5
<211> 69
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(69)

<400> 5
acc cac tcc tgt gtg gac ctg gat gac aag ggc tgc ccc gcc gag cag 48
Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly Cys Pro Ala Glu Gln
1 5 10 15

aga gcc agc cct ctg acg tcc 69
Arg Ala Ser Pro Leu Thr Ser
20

<210> 6
<211> 23
<212> PRT
<213> Homo sapiens

<400> 6
Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly Cys Pro Ala Glu Gln
1 5 10 15

Arg Ala Ser Pro Leu Thr Ser
20

<210> 7
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(21)

<400> 7
acc cac tcc cct ctg acg tcc 21

Thr His Ser Pro Leu Thr Ser
1 5

<210> 8
<211> 7
<212> PRT
<213> Homo sapiens

<400> 8
Thr His Ser Pro Leu Thr Ser
1 5

<210> 9
<211> 42
<212> PRT
<213> Homo sapiens

<400> 9
Cys Gln Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp
1 5 10 15

Asp Lys Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile
20 25 30

Ile Ser Ala Val Val Gly Ile Leu Leu Val
35 40

<210> 10
<211> 42
<212> PRT
<213> Homo sapiens

<400> 10
Cys Gln Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp
1 5 10 15

Asp Lys Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile
20 25 30

Ile Ser Ala Val Glu Gly Ile Leu Leu Val
35 40

<210> 11
<211> 30
<212> PRT
<213> Homo sapiens

<400> 11
Cys Gln Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Ala Ser Pro
1 5 10 15

Leu Thr Ser Ile Ile Ser Ala Val Val Gly Ile Leu Leu Val
20 25 30

<210> 12
<211> 26
<212> PRT
<213> Homo sapiens

<400> 12
Cys Gln Pro Cys Pro Ile Asn Cys Thr His Ser Pro Leu Thr Ser Ile
1 5 10 15

```
Ile Ser Ala Val Val Gly Ile Leu Leu Val
            20                  25
```

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 99/00912

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/71 C12N15/85 C12N5/10 A61K38/17 A61K48/00 A61K39/395 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SIEGEL P M ET AL: "Novel activating mutations in the neu proto-oncogene involved in induction of mammary tumors." MOLECULAR AND CELLULAR BIOLOGY, vol. 14, no. 11, November 1994 (1994-11), pages 7068-7077, XP000867018 cited in the application	1-16, 18-21
Y	figures 1,5 and 6 table 1 page 7073, right-hand column -page 7074, left-hand column	15,16, 18,19

-/--

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 February 2000	17/03/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ALCONADA RODRIG.., A

6

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KWONG, K. Y. ET AL: "Identification of a novel alternative splicing form of human HER2 / neu proto-oncogene." PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL MEETING, (MARCH, 1998) VOL. 39, PP. 205. MEETING INFO.: 89TH ANNUAL MEETING OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH NEW ORLEANS, LOUISIANA, USA MARCH 28-APRIL 1, 1998 AMERICAN, XP000866525	5,9-13, 15,16, 18-21
Y	abstract	15,16, 18,19
Y	SHEPARD H M ET AL: "MONOCLONAL ANTIBODY THERAPY OF HUMAN CANCER: TAKING THE HER2 PROTOONCOGENE TO THE CLINIC" JOURNAL OF CLINICAL IMMUNOLOGY,US,PLENUM PUBLISHING CO, vol. 11, no. 3, 1 May 1991 (1991-05-01), page 117-127 XP000560916 ISSN: 0271-9142 page 119-125 figure 2; tables I-IV	18,19
Y	COLOMER R ET AL: "erbB-2 antisense oligonucleotides inhibit the proliferation of breast carcinoma cells with erbB-2 oncogene amplification" BRITISH JOURNAL OF CANCER, vol. 70, no. 5, November 1994 (1994-11), pages 819-825, XP000867330 figures 1-3 see results section	15,16, 18,19
P,X	KWONG K Y ET AL: "A novel splice variant of HER2 with increased transformation activity." MOLECULAR CARCINOGENESIS , vol. 23, no. 2, October 1998 (1998-10), pages 62-68, XP000867021 figures 1,2,4 table 1 page 64	1-16, 18-21

-/--

6

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	SIEGEL P M ET AL: "Elevated expression of activated forms of Neu / ErbB -2 and ErbB -3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer." EMBO JOURNAL, vol. 18, no. 8, 15 April 1999 (1999-04-15), pages 2149-2164, XP002131033 page 2154, right-hand column -page 2157, left-hand column figures 9-11 table IV	1-16, 18-21

6

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 12-16,18 and 19, as far as concerning an in vivo method are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 17
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
See FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17

Present claim 17 relates to an extremely large number of possible methods. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found only for the methods involving the use of antibodies against the SPLICE erbB-2 protein or the use of erbB-2 antisense oligonucleotide. In the present case, the claim so lacks support, and the application so lacks disclosure, that a meaningful search over part of the claimed scope is impossible. Consequently, the search has not been carried out for that part of the claim that does not refer to the use of anti-SPLICE erbB-2 antibodies or antisense oligonucleotides complementary to the sequence of the polynucleotide encoding the SPLICE erbB-2.

The applicant's attention is drawn to the fact that claims, or part of claims, relating to inventions in respect of who no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

(Remark: Present claims 3 and 4 refer, respectively, to a nucleic acid sequence encoding the SPLICE erbB-2 protein having the amino acid sequence according to SEQ ID NO:2 and to the nucleic acid sequence according to SEQ ID NO.1 encoding a SPLICE erbB-2 protein. However, the sequences provided in the description of the application (both in the computer-readable form and in figures 1 and 2) correspond to the wild-type human erbB-2 polynucleotide (SEQ ID NO.1) or polypeptide (SEQ ID NO:2). Therefore, the search for said claims was carried out considering the sequence of the alternate splice human erbB-2 provided in figure 3B.)